

Half a World: Regional inequality in five great federations

Branko Milanovic¹

World Bank and Carnegie Endowment for International Peace

The paper studies regional (spatial) inequality in five most populous countries in the world: China, India, the United States of America, Indonesia and Brazil in the period 1980-2000. They are all federations or quasi-federations composed of entities with substantial economic autonomy. Two types of regional inequalities are considered: Concept 1 inequality which is inequality between mean incomes (GDPs per capita) of states/provinces and Concept 2 inequality which is inequality between population-weighted regional mean incomes. The first inequality speaks to the issue of regional convergence, the second, to the issue of overall inequality as perceived by citizens within a nation. All three Asian countries, show rising inequality in terms of both concepts in the decade of the 1990's. Divergence in income outcomes is particularly noticeable for the most populous states/provinces in India and China. United States, where regional inequality is the least, shows further convergence. Brazil, with the highest level of regional inequality, displays no trend. A regression analysis fails to establish robust association between the usual macro variables and the two types of regional inequality.

Key words: China, India, USA, Brazil, Indonesia, regional inequality, world inequality
JEL classification: R12, I3, O57

World Bank Policy Research Working Paper 3699, September 2005

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.

¹ I am grateful to Tamar Manuelyan Atinc, Roberto Zagha, Bala Bhaskar Naidu and Bert Hofman for providing most of the regional GDP data used in this paper. Gouthami Padam was an excellent research assistant, and I am grateful to her. I am grateful to Isabelle Bensidoun, Guillaume Daudin, Bert Hofman, Ravi Kanbur, Bert Keidel, Christian Morrisson, Mansoob Murshed, Martin Ravallion, Erik Thorbecke, Xiaobo Zhang, an anonymous referee, and participants of a conference on inequality in China at Cornell University and at Brookings Institution for very useful comments. Work was financed by a grant from McArthur Foundation and by the World Bank Research Grant RPO 85725. Email: bmilanovic@worldbank.org.

1. Introduction

Global inequality is strongly influenced by what happens to populous countries, both to their average income levels, and inequality within each country. If we look at the former, fast growth rates of China and more recently of India, by reducing their relative income distance from the United States and other rich countries, have lowered global inequality. Fast (average) growth in these nations is the most important element contributing both to lower global inequality and to lower global poverty. At the same time however within these nations, there are increasing cleavages: some provinces grow faster than others, so mean income differences between the provinces increase. In addition, the process of growth is accompanied in both China and India by increasing inequality between individuals in the country as well as within most provinces and states (see for example, Jha (2004) for India; Chen and Wang (2001) for China).

The objective of this paper is to explore the subnational changes in the five most populous countries in the world: China, India, United States, Indonesia, and Brazil. We examine these changes for three reasons. First, as already mentioned, subnational changes in populous countries are important because they are a significant determinants of global inequality. The second reason lies in what subnational changes imply for national unity. One of the concerns of the Chinese leadership, for example, is how rising income differences between the prosperous coastal provinces and the less dynamic North East may affect China's political unity (see Renard, 2002).² The third reason to look at the subnational level is that it provides us with some lessons regarding what we might expect at the global level were most of the obstacles to free circulation of labor, capital and goods to be lifted. For obviously although these obstacles still exist in a number of countries even between their constituent units, they are dramatically lower than between nations. For example, shipment of goods between different Indian states is, in many

² In June 2004, nine Chinese provinces (Yunnan, Guizhou, Guangdong, Fujian, Jiangxi, Hunan, Guangxi, Hainan and Sichuan) and the special administrative regions of Hong Kong and Macau created a Pan Pearl River Regional Cooperation Council whose objective is to eliminate nontariff barriers between the provinces and to coordinate provincial economic policies. One of the objectives was to speed up growth in order to catch up with Shanghai and the booming Yangtze River basin (see International Herald Tribune, "Chinese provinces form a trading bloc", June 2, 2004, p. 17).

cases, subject to fees or border check but this is a far cry from the obstacles to international trade. Similarly, labor mobility in China is formally restricted as people need to have a permit in order to live in large cities but these rules are most often ignored, and even if they were fully implemented, the rules are much less stringent than the rules governing circulation of labor between nations. It is thus important to find out whether a generally unhindered movement of goods, labor and capital leads provinces toward convergence in their mean incomes or not.

A comparison between subnational inequalities in these five countries is meaningful despite the fact that the number and size of regional units vary. If states and provinces were based on randomly drawn border lines, it could be indeed argued that greater regional inequality in one country compared to another is simply an artifact, possibly due to the way such regional units are defined. However, in most or all cases with which we deal here, provinces/states possess both a significant amount of autonomy and specificity. Inequality between them thus does have a bearing on the issues of horizontal equity, and particularly so where states differ in their religious and ethnic composition.

The structure of the paper is as follows. In Section 2, we present some data that show the importance of the five countries in total world income, population and global inequality. In Section 3, we move to the subnational level of analysis and present a brief review of the literature dealing with the issue of regional inequality in the five countries. In Sections 4, we look at the evolution of the differences between mean incomes of the regions during the last twenty years in each of the five nations.³ It is here that we address the issues of income convergence or divergence within each country. In Section 5, we look at the differences in mean regional incomes but now with each region weighted by its population share. Once we weigh them by their populations, we come closer to

³ We use the term “region” to indicate a subnational unit whether its exact appellation is “state” or “province” or “republic”; the term “nation” or “country” is used to designate the subjects of the international law, that is the five countries included in the analysis here.

explaining the actual “feeling” of inequality within each nation particularly when differences in mean regional incomes coincide with other horizontal cleavages like ethnicity or religion. We also come closer to explaining the overall national inequality between individuals because the only part which is not included now is the part of inequality—at times substantial, it is true—which is due to inter-personal income differences within each region. In Section 6, we try to identify the policies or factors at national level which might explain the process of regional convergence or divergence; in other words, we try to find some regularities explaining why regions are converging or diverging. Section 7 presents the conclusions.

2. The importance of the five countries in the world

Table 1 shows the share of the five most populous countries in world population and income in 1980, 1990 and 2000. Their share in world population has throughout remained at just above one-half. However, their share in world output has expanded from a bit over one-third to 44 percent.

Table 1.Share of the five largest countries in world population and world income (in %)

	1980	1990	2000
Share in world population	51.7	50.6	50.8
Share in world PPP output	34.3	36.6	44.0

The increase in the five countries’ share in world income was driven by the growth (compared to world mean income) of all countries save Brazil (Table 2). The most extraordinary was of course China’s growth. In 1980, China’s GDP per capita was just under 17 percent of the world average. In 2000, it was in excess of 60 percent. The United States likewise grew faster than world mean income: its GDP per capita was 4.3 times greater than the world mean in 1980, while in 2000, it was almost 4.6 times as high.

Table 2. The five large countries' GDPs per capita compared to the world mean

	1980	1990	2000	Change (1980-2000)
India	0.164	0.206	0.246	+0.082
China	0.169	0.307	0.602	+0.433
Indonesia	0.364	0.495	0.527	+0.163
Brazil	1.029	0.869	0.822	-0.207
United States	4.308	4.470	4.579	+0.271

Note: All calculations done for GDP per capita expressed in 1995 international (PPP) dollars. Source: World Bank.

What is the importance of these countries for global inequality? Being the most populous countries in the world, they will obviously influence, sometimes decisively, global inequality. However when we address the issues of global inequality we have to be very careful in what exactly we are measuring. The first concept of inequality (dubbed Concept 1 by Milanovic, 2005)) measures differences in mean incomes between countries. No weighting is involved here and each country counts the same. This concept is important if we are interested whether there is unconditional divergence or convergence of mean (countries') incomes in the world, but there our five countries count as much as any other five countries in the world.

Concept 2 inequality is different. It likewise takes mean national incomes but weighs them by countries' populations. Now, of course, our five countries will matter a lot because they account for almost one-half of world income and population. But note that when we use Concept 2 inequality, we disregard the entire inequality due to the differences in personal incomes within nations, that is we assume that each individual has the mean income of his/her country. Concept 3 inequality (inequality between all individuals in the world) is accordingly underestimated by Concept 2 inequality.⁴

Table 3 shows that the differences in mean incomes (GDPs per capita) between the five countries (that is, disregarding their interactions with other countries) account for between a fifth and a third of total Concept 2 inequality in the world (depending on

⁴ However about $\frac{3}{4}$ of Concept 3 inequality in the world is accounted for by the differences in countries' mean incomes (Milanovic, 2002 and 2005).

whether we use Gini or Theil index). It is important to explain how this calculations is done. Consider the Gini coefficient. In the case of Concept 2 inequality, it is equal to

$$\frac{1}{\mu} \sum_i^n \sum_{j>i}^n (y_j - y_i) p_i p_j \quad (1)$$

where y_i = GDP per capita of i -th country (and $y_j > y_i$), p_i = population share of i -th country, and μ = mean world income. Thus, Gini Concept 2 for the world is calculated by taking the absolute differences in mean incomes between all countries in the world, normalizing them by the world mean income and weighing by the population shares. To calculate the part due to the five countries we “extract” only those individual inter-country terms (ICT), that is $\frac{1}{\mu} (y_j - y_i) p_i p_j$, where both i and j belong to “our” five

countries. Thus, for example, in 1980, the US GDP per capita was \$PPP 21,755 and China’s \$PPP 852. The difference between the two amounted to \$PPP 20,900. In terms of world mean income (which was then \$5,050), the difference was 4.13 times. This is in turn weighted by US and Chinese population shares (0.054 and 0.234 respectively), and the obtained value 5.3 Gini points represents the contribution of the difference in mean incomes between the United States and China to Concept 2 world Gini. The same calculation is repeated for all five countries (there is in total 10 such calculations). The sum gives the total amount of the direct contribution due to the interaction between the five countries. As Table 3 shows, it did not change much in the last twenty years: it oscillated around 11 Gini points throughout.

Note however that these calculations leave out all the similar interactions between each of our five countries and the rest of the world (say, China vs. Germany, or US vs. Nigeria, or India vs. Senegal etc.) which also contribute to world inequality. If we add all these contributions, we notice that they have declined from almost 17 Gini points in 1980 to 14 Gini points in 2000. The decline was almost entirely due to China’s fast growth that diminished the distance, and hence the inequality terms, between herself and the rest of the (rich) world. Overall the five largest countries contribute around one-half of total global Concept 2 inequality as measured by the Gini coefficient.

Differently, we can measure inequality by using the Theil index. The Theil index for Concept 2 inequality can be written:

$$\sum_{i=1}^n p_i \frac{y_i}{\mu} \ln \frac{y_i}{\mu} \quad (2)$$

Here we take the same approach: we “extract” only the $p_i \frac{y_i}{\mu} \ln \frac{y_i}{\mu}$ terms belonging to our five countries. The contribution of the US is obtained (again using the 1980 data) as the product of its population share (0.054) with its income relative to the world mean (20,900/5050=4.308) times the log of the last amount. To get the total contribution of the five countries, we simply add such amounts.⁵ Their contribution to the global Theil-measured Concept 2 inequality varies between one-third in 1980 and 37 percent in 2000 (see Table 3).

Table 3. Direct and total contribution of the five largest countries to Concept 2 global inequality

	1980	1990	2000
Gini			
Concept 2 Gini (world)	54.1	52.9	50.2
Interaction between the five countries (Gini points)	11.6	10.8	11.3
Interaction of the five with other countries (Gini points)	16.9	15.6	14.0
Total percentage contribution	53	50	50
Theil entropy			
Concept 2 Theil (world)	63.3	59.8	52.3
Contribution of five largest countries (Theil points)	21.0	17.8	19.4
Percentage contribution	33	30	37

Note: Concept 2 inequality is inequality between mean countries' GDPs per capita weighted by countries' populations. Both Gini and Theil expressed in percentages.

⁵ Unlike the Gini coefficient which is based on bilateral comparisons of income between all countries and allows us to distinguish the contribution of each and every pair of countries, Theil index just sums individual countries' contributions. In other words, it cannot distinguish between the contribution due to the interaction between the five countries themselves, and the five countries and the rest of the world.

3. A brief review of regional (within-country) inequality studies⁶

The issues we address here—Concept 1 and Concept 2 inequality *within* nations—have been, in slightly different contexts, addressed before. This was done in two contexts. The first is the issue of regional inequality within countries. There are two views in the literature that are often juxtaposed. The first is due to Williamson (1965) who argued that in the early stages of economic development, regional inequality would tend to rise as growth occurs in discrete locales, but that later inequalities will decline as equilibrating forces such as better infrastructure, technological diffusion, decreasing returns to capital in richer and high-wage areas, diseconomies of agglomeration etc. become stronger. Thus, regional inequality is expected to follow an inverted U shape as income level grows. Williamson's reasoning is closely related to the idea of the Kuznets curve where similar development although not in spatial terms produces first an increase and then a decline in inequality. It is also based on the neoclassical (Solow-type) assumptions which include decreasing marginal returns. A different view has been proposed more recently within the context of the new economic geography school (Krugman and Venables, 1995) and endogenous growth (Romer 1986; see also a recent review by Easterly and Levine 2002). There the argument is that that increasing returns to scale and advantages of agglomeration of capital and knowledge will tend to perpetuate, or even increase, spatial inequalities. Yet in Krugman and Venables (1995), decreasing transportation costs may play an offsetting role: assume that transportation costs are zero, then the advantage of cheap labor in less developed countries (or regions) will, to some extent, tend to offset the advantages of increasing returns to scale.

The key difference between these two approaches seems not to lie in their view of the short-run developments, where they all, including the earlier development theories such as Myrdal's (1957), Rosenstein-Rodan's (1943), Hirschmann (1958) or Perroux (1970; 1988), seem to agree that growth is disequalizing, but in their view of the long-run developments where either traditional neoclassical assumptions dominate—rendering

⁶ Since we deal with *regional* inequality, we do not review studies of the most common Concept 3 (inter-personal) inequality.

growth ultimately equalizing in spatial terms as well—or where such assumptions are rejected or made less potent thus weakening the forces which make for spatial equality.⁷ Recently, the short- and long-run aspects have been combined in a paper by Petrakos, Rodriguez-Pose and Roviolis (2003), which looks at the regional inequalities within the European Union (with several regions defined within each country). The authors find that the short-term effects of growth are disequalizing in the sense that higher growth rate tends to increase regional inequality (controlling for all other country-relevant attributes), while higher income *level* is associated with lower regional inequality. The authors interpret the second finding as implying the long-run equilibrating effects of growth along the lines of the Solow and Williamson models. Their measure of regional inequality, as in several other papers (e.g. Akita and Kataoka (2003) regarding Japan; Akita and Kawamura (1992) regarding Indonesia and China; Bhalla, Yao and Zhang (2003) and Kanbur and Zhang (2003) for China) is the population-weighted coefficient of variation or population weighted Theil index. This is what we called Concept 2 inequality, and the justification for using Concept 2 (rather than Concept 1) inequality is that it reflects regional inequality as experienced by an “average” person; in other words, regional divergence which may be due to a few sparsely populated regions’ either very fast or very slow growth is rather irrelevant for the actual feeling of spatial (horizontal) inequality as experienced by the people in the country.

But the issue of regional inequality—using Concept 1 inequality—has also recently received quite a lot of prominence due to the popularity enjoyed by the so-called growth convergence literature. While the convergence issues have originally been defined and studied at the level of countries (that is, convergence of national economies), they have recently been studied at the level of subnational regions. For the countries included here, such papers are Zhang, Liu and Yao (2001) for China, Azzoni (2001) for Brazil, Ram (1992) for the United States, and Jha (2004) and Dreze and Srinivasan (2000) for India.⁸ The rationale for the interest in Concept 1 inequality is very different from the interest in

⁷ There is an obvious link between these views, as couched in terms of regional developments within individual countries and regional developments in the world as in Krugman and Venables (1995) or Krugman (1991).

Concept 2 inequality. The issue of convergence of (unweighted) regions within a country, or (unweighted) countries in the world, addresses the problem of whether the same or similar economic policies produce similar results or not. Consider the example of a country which has a single national economic policy, that is where there is no significant regional freedom of economic policy-making. Suppose that Concept 2 regional inequality is decreasing. But if we still find that Concept 1 inequality is increasing, it raises the interesting question of what characteristics enjoyed by some regions are responsible for their not catching up (or for their growing too fast). Thus both Concept 1 and Concept 2 inequality are of interest.

We shall now briefly review some of key (representative) regional inequality studies that deal with the five great federations included here.⁹ The studies of China are the most numerous. There are two reasons for this. First, the extremely fast growth of Chinese economy over the last quarter century has been associated with increasing regional inequalities. This has obvious implications both for political stability and for economic theory, that is for figuring out why and how certain regions grow and others don't, and whether the dominant feature of China's inequality is rural vs. urban inequality, or inter-provincial inequality. The consensus seems to be that it is the former.¹⁰ For example, Bhalla, Yao and Zhang (2003) calculate inequality in per capita consumption across provincial and urban-rural partitions (that is, they use data for mean annual incomes for rural and urban areas for each province, that is 28 provinces times 2 = 56 observations¹¹) and find that in 1995, more than $\frac{3}{4}$ of thus calculated Theil Concept 2 inequality is accounted for by the rural-urban split.¹² This is a result similar to the one obtained by Kanbur and Zhang (2003) who find that the same urban-rural split (that is, the difference

⁸ For other countries, see Goerlich and Mas (2001) for Spanish provinces, Yemtsov (2002) for the subjects of the Russian Federation.

⁹ We use the term "federation" in a technical sense, to indicate that the constituent parts do have some power of economic decision-making and represent meaningful entities in historic, ethnic or religious sense. Not all of the five countries studied here are federations in a juridic sense of the word (see also Annex 1).

¹⁰ When it comes to inter-provincial inequality, it seems to be more the case of "club" inequality, that is of three clubs (East, West, Center) diverging from each other (see Yao and Zhang 2001).

¹¹ More exactly, they have the data on mean per capita consumption of peasants and non-peasants (by province) as obtained from Chinese household surveys. They interpret peasant consumption to be rural, and likewise non-peasant to be equivalent to urban (see Bhalla et al., p. 945).

¹² It is notable that the share of the urban-rural difference in total appears constantly high, that is between 70 and 80 percent, from 1952 to the end of century (see Bhalla et al (2003, Table 2, p. 947)).

between mean urban and mean rural income by province) explains 56 percent of Concept 2 inequality (calculated from the same 56 observations).¹³ The results of these two studies and a few others are compared in Annex 2.

The second reason lies in the lack of individual-level data on income inequality in China, that is lack of data on Concept 3 inequality.¹⁴ Concept 2 regional inequality can be used, if our partitioning is sufficiently fine, to approximate the evolution in Concept 3 inequality. In other words, if we think that most of inequality is spatial, and use a very fine partition (divide the country in meaningful but also numerous regional units) then thus calculated Concept 2 inequality should approximate, if not necessarily the level then the evolution, of Concept 3 inequality. To see this consider that, at the extreme, every individual can be treated as a “region”: then Concept 2 inequality collapses into Concept 3 inequality. This was, for example, the approach underlying Kanbur and Zhang (2003) paper on regional inequality in China. As in the Bhalla et al. (2003), Kanbur and Zhang divide China into 28 provinces and each of the provinces into its rural and urban areas. They have the data for mean incomes for each of thus defined 56 regions for the period 1952-2000. They calculate Concept 2 inequality from these means, find that the rural-urban split accounts for the bulk of total inequality (much more than the inland-coastal split)¹⁵ but use the Concept 2 inequality as a proxy for the Concept 3 inequality. Then they try to relate changes in Concept 2 inequality levels over the last 50 years to various policy episodes (Great Leap Forward, Cultural Revolution, agricultural liberalization, urban and industrial liberalization etc.) They find that the Concept 2 inequality, with this relatively fine partition, amounts in 2000 to a Gini of 37.2 which, of course, sets a rather high lower bound to total personal income inequality in China.¹⁶

¹³ Why these two results are not the same is puzzling. A comparison of the results of these two studies and a few others can be found in Annex 2 of the Internet version of this paper available at <http://econpapers.hhs.se/RAS/pmi44.htm>

¹⁴ Of course, there are many studies of inter-personal inequality in China. They are however all approximations based on the published group data from national surveys since Chinese authorities have been unwilling to share micro (individual-level) data.

¹⁵ Although after 1993, there is a rapid increase in the within-urban and within-rural components indicating that there are widening income differentiations *within* urban and *within* rural areas as well. (“Within” in this context means “between mean incomes of rural (or urban) provincial incomes.”)

¹⁶ When we use regional GDP per capita as welfare indicators (as here) to derive Concept 2 inequality, its value cannot be fully taken as the lower bound of Concept 3 inequality because of likely transfers between

As for other countries, Jha (2004), in a study of India's inequality over the last fifty years looks at the issue of Concept 1 convergence between the states and concludes that divergence has been more common and has accelerated since the reforms in the early 1990's. Ram (1992), and Barro and Sala-i-Martin (1992) have done similar analysis for the United States. Ram (1992) finds a steadily decreasing inter-state inequality from 1950 to 1980 and an increase in the next decade. For Brazil, Azzoni (2001, p. 137) shows decreasing Concept 1 inequality throughout the 1950's and 1960's and a stable one in the last twenty years.

As this brief review shows, regional inequality studies fall into three categories that closely match our three concepts of inequality. Many of them, in the wake of the convergence literature, deal with Concept 1 inequality. Others, perhaps equally numerous, deal with Concept 2 inequality—regional or horizontal inequality as actually “experienced” by the population. Finally, some use regional partition (Concept 2 inequality) as a proxy for Concept 3 inequality.

It should be noted that the work on regional inequality is not facilitated by the absence of an accepted or clear terminology. The results are often impenetrable because the same term, say “regional inequality” may be used to represent Concept 1 or Concept 2 inequality. Even the term “region” is sometimes used for the smallest units (say, states in a country) and, perhaps in the same paper, for the agglomeration of several such units into a larger whole which is still not national level (thus, for example, authors often write of China's three regions: East, Center and West, *and* of China's regions, meaning in the latter case provinces). As a consequence, “regional inequality” might mean either one of the four combinations: Concept 1 or Concept 2 inequality, or inequality between the provinces or inequality between agglomerations of provinces, that is larger “regions.” In addition, the share of the between (inter-regional) component is quite different—even if

the regions. This is different from the analysis on the global level when redistributive transfers between the countries are minimal. I owe this point to Christian Morrisson.

the partitions are the same—if we look at the “between” share in total interpersonal (Concept 3) inequality or in Concept 2 inequality.¹⁷

Even the measures of inequality are often opaque. While many authors use Theil indexes because of their decomposition properties, it is not always clear if Theil (1) also known as Theil entropy measure, or Theil (0) also known as mean log deviation index is used. While the use of one or another Theil index does not have an impact on results within each individual study (since the measures move almost always the same way), it does render difficult comparison of the absolute values of inequality measures from several different studies.¹⁸ We have tried to be as clear as possible in our terminology here and to the risk of being repetitive will use “Concept 1” and “Concept 2” inequality often in order to avoid possible misunderstanding. Similarly, we shall try to be clear about the “partition” we use: states (provinces) adding up to a nation, or states (provinces) adding up to regions which then add up to a nation.

4. The five federations: descriptive statistics and Concept 1 inequality

The analysis of each individual country’s inequality is conducted on the basis of regional GDPs per capita expressed in nominal prices.¹⁹ Clearly, to the extent that there are price differences within a country and price levels are higher in richer regions, such inequality statistics will tend to overestimate intra-regional inequality.

In order to allow for comparison between regions belonging to different countries, all regional GDPs are also expressed in real terms. First, they are expressed in real units of the domestic currency and then converted into the 1995 international dollars (\$PPP). The typical pattern of conversion is the one for China as expressed in equation (1) where Y denotes provincial GDP per capita, r = real provincial growth rate, DD = country-wide

¹⁷ For example, in a very detailed summary of the results of the “between” component in different country studies given in Sharrocks and Wan (2004, Table 1A), total inequality against which the “between” component is measured is sometimes Concept 2 and sometimes Concept 3 inequality.

¹⁸ Yet another problem seems to be that some authors use natural logs and others use logs with the base of 10. Again, this does not matter for individual studies, but does matter for comparative work.

¹⁹ Except in Indonesia for which we have only real provincial incomes.

deflator, and PPP = PPP-equivalent dollar exchange rate. We have, from the Chinese statistics, real annual provincial growth rates expressed in 1978 all-China prices. This means that real income of region i , country j (=China here) and year t (subsequent to 1978) is obtained by applying real growth rates to the nominal 1978 regional income. Note that the application of real growth rates from region to region implies that differences in annual inflation between the regions are accounted for. This amount, that is provincial income in year t expressed in 1978 prices (the denominator in equation 3) is then, using the country-wide deflator between 1978 and 1995 ($DD_{j,78,95}$), converted into all-China 1995 prices.²⁰ This is finally converted into international dollars by using the 1995 purchasing power exchange rate of the country (obtained from the World Bank data).

$$Y_{i,j,t,95,\$} = \frac{Y_{i,j,t-1,78,d} * (1 + r_{i,j,t})}{DD_{j(78,95)}} * PPP_{95} \quad (3)$$

where the subscript d denotes domestic constant prices, and subscript $\$$ constant international prices.²¹

A slightly different approach is used for US regions. Nominal incomes (year t expressed in dollars of the same year) are converted using the overall country's deflator (DD_j) which converts t year's dollars into the 1995 dollars. This implies that price differences across US states are non-existent.²² Since the US prices are, by construction, made to equal to international prices, the PPP convertor is equal to 1 (equation 4)

$$Y_{i,j,t,95,\$} = \frac{Y_{i,j,t,t,\$}}{DD_{j(t,95)}} * 1 \quad (4)$$

Table 4 shows the descriptive statistics and our data coverage of the five countries. More detailed information regarding countries' administrative structure is given in Annex 1. Figure 1 shows Concept 1 inequality for the five countries calculated using Gini and Theil indexes.

²⁰ Note that the deflator has only country (j) subscript.

²¹ A special caveat must be made regarding Chinese national account data: see the discussion in Annex 1.

²² There is also no state-level CPI in the United States.

We see that the unweighted level of inter-regional inequality is the highest in China, next come Brazil and Indonesia,²³ followed by India and the lowest level of inter-regional inequality is in the United States, despite the fact that the US has the highest number of states and that there is a presumption that Concept 1 inequality will increase with greater number of partitions (see Proposition 1 in Shorrocks and Wan, 2004). As we can see from Figure 1, the Concept 1 Gini is about 30 for China and Brazil, 25 for Indonesia, less than 20 for India, while for the United States is it around 10. The ranking of the countries is the more or less same if we use Theil index—except that in the year 2000 Indonesia displays higher inequality than Brazil.

But the figures also reveal very different time patterns in the evolution of Concept 1 inequality. China which has the highest levels of inequality displays also significant reductions throughout the 1980's. The decline in regional inequality is particularly impressive if we look at the Theil index which puts greater emphasis on what happens at the tails of the distribution. The period of regional equalization however comes to a close around 1990 and is reversed afterwards. On the contrary, Indonesia and India show increasing regional inequality throughout the entire period; Brazil, no noticeable trend. Finally, the United States which almost throughout has the lowest level of inter-regional differences, shows a burst of increased inequality in the early 1980s, followed by a sustained if weak reduction since.

When we look at the evolution of inter-regional inequality within a country, we need to distinguish between inter-regional inequality calculated in nominal and real prices. As discussed earlier, the former is simply a comparison of regional nominal GDPs per capita; the latter is calculated based on GDP per capita and a price structure of a given year which is then “augmented” by annual real growth rates. The results presented in Figure 1 are based on nominal magnitudes with the exception of Indonesia where we

²³ In accordance with the rest of the literature (see Akita and Kawamura, 2002, and Akita and Lukman, 1995) regional inequality statistics for Indonesia are based only on non-oil and gas part of the GDP. Inequality is significantly greater if we include oil and gas but on the other hand most of these revenues are not retained at the provincial level but are appropriated by the Center and then redistributed. Hence such inequality would overestimate actual regional inequality in the standard of living of the population.

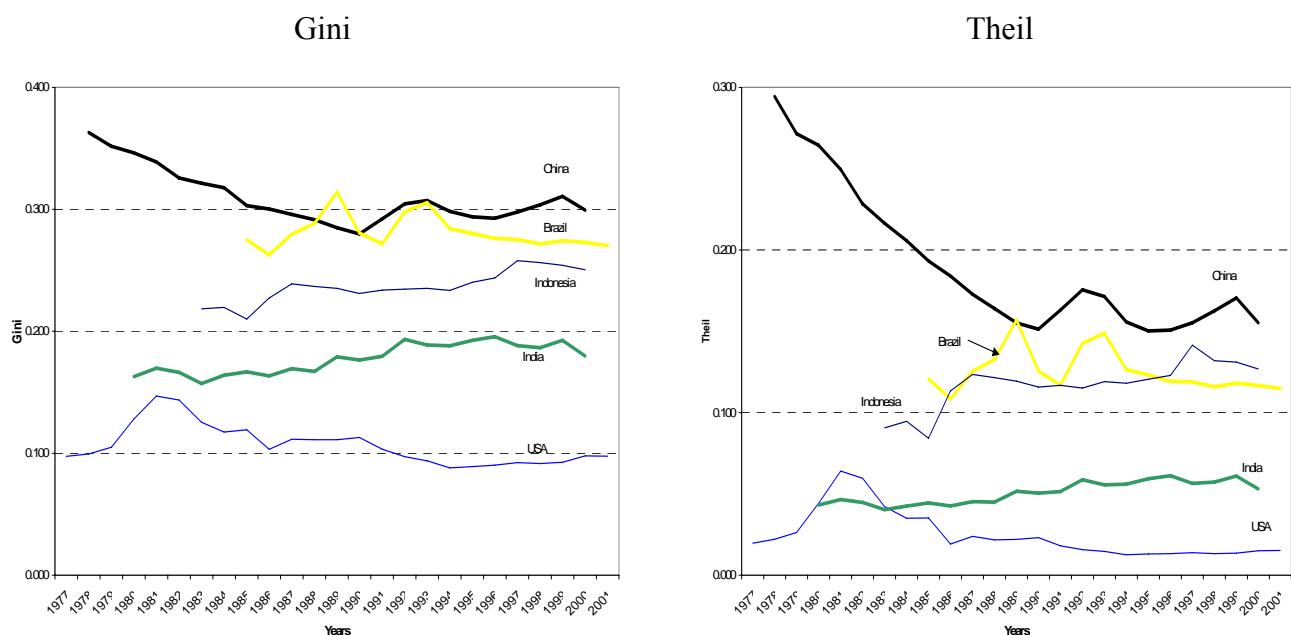
had access to real data only. However, in China and India nominal and real inequality measures do not move in parallel.

Table 4. Regions in the five countries: descriptive statistics

	China	India	USA	Brazil	Indonesia
Period included	1978-2001	1980-2000	1977-2001	1985-2001	1983-2001
Welfare indicator	GDP per capita	GDP per capita	GDP per capita	GDP per capita	GDP per capita
Constant prices (year)	1978	1980-81	1995	1985	1983
Number of regions included in the study	29	14	50	26	26
Total number of regions in the country	30	25	50	26	26
Not included in the study (approx. population in m)	Hong Kong, Macau, Taiwan (China) 29 million	All Union territories (Delhi etc.), 11 million	District of Columbia, Pacific and Caribbean possessions, 5 million 2/	Federal District, 2 million	--- 3/
Total population (million; in 2000)	1,271	1,033	283	172	213
Population coverage (in percent; 2000) 1/	99	92	100	99	100
Most populous region (million, 2000)	114.2 (Sichuan)	170.6 (Uttar Pradesh)	34.0 (California)	37.1 (Sao Paulo)	35.7 (West Java)
Least populous region (million, 2000)	5.6 (Ningxia)	20.6 (Haryana)	0.5 (Wyoming)	0.3 (Roraima)	1.2 (Maluku)
Richest region (year 2000; in tho. 1995 international dollars)	21.7 (Shanghai)	3.1 (Maharashtra)	42.0 (Connecticut)	9.1 (Sao Paulo)	13.7 (East Kalimantan)
Poorest region (year 2000; in tho. 1995 international dollars)	1.6 (Guizhou)	0.7 (Bihar)	20.0 (West Virginia)	1.3 (Maranhao)	1.1 (East Nusa Tenggara)
Ratio: richest-poorest region (2000)	13.6	4.4	2.1	7.0	12.5
Median region by income (year 2000; in tho. 1995 international dollars)	3.3 (Anhui)	1.7 (Kerala)	28.8 (Pennsylvania)	3.1 (Pernambuco)	2.4 (East Java)
GDP per capita (year 2000; in tho. 1995 international dollars)	4.1	1.7	31.5	5.7	3.6

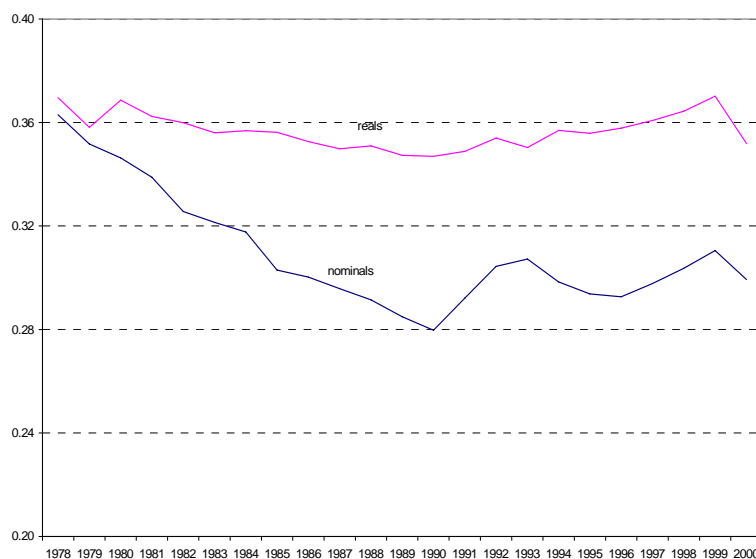
Note: For the purposes of comparison, GDP per capita is always expressed in 1995 international dollars. No region-specific PPP rates are used (e.g. all regions of a given country have their GDP per capita converted into \$PPP using a single exchange rate). 1/ Population in regions included in the study divided by total population as defined here. 2/ Pacific and Caribbean possessions include 12 territories only 5 of which (Puerto Rico, Guam, US Virgin Islands, American Samoa and Mariana Islands) have resident population apart from the military. Their total population is 4.4 million (see <http://www.infoplease.com/ipa/A0108295.html>). 3/ East Timor was included in the Indonesian data until independence in 1999. For consistency, we have excluded it throughout.

Figure 1. Concept 1 (unweighted inter-regional inequality) in the five countries



Note: Concept 1 inequality is calculated in nominal terms (except for Indonesia where we have real data only). GDP for Indonesia excludes oil and gas data.

Figure 2. China: Concept 1 inequality calculated in nominal and real terms (1978 prices)

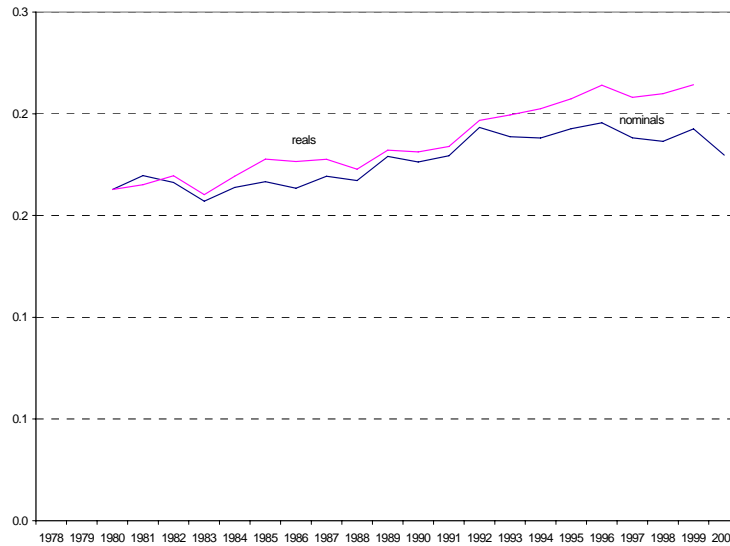


Note: Provinces of Qinghai and Hainan (populations of respectively 5 and 7 million) are not included in the calculations based on real GDP per capita due to the lack of data. They are included in the nominal GDP data series though.

Figure 2 shows very different evolutions of the two inequalities in the case of China. If we look at nominal differences in GDP per capita, we note a strong reduction in inter-provincial inequality up to 1990, and a slightly increased level since. The decline in nominal inequalities (which is even more dramatic if we look at the Theil than at the Gini measure and thus implies that the catch up has been particularly strong for the poor provinces) coincides with the period of agricultural liberalization. Since 1990, approximately around the time when China enters more substantial liberalization in the urban sector, this trend is reversed. However if we look at the inter-provincial inequality in real terms, we see that it remained around Gini of 35 throughout the entire period. The implication of these different movements in real and nominal Gini and Theil is that there has been a price catch up of the poorer provinces (or that the output mix produced in poor provinces has moved toward products whose prices have been rising more than the average). In other words, if in 1978, the poorer provinces had a lower price level,²⁴ then while their real growth rates did not systematically differ from those of the rich provinces, their relative price level must have risen in order to observe a decline in nominal inter-provincial inequality. After 1990, the two measures move the same way indicating that the price catch-up has stopped.

²⁴ This is a very sensible supposition although we cannot prove it since Chinese statistics publish only growth rates of the provincial CPIs but do not provide provincial price levels.

Figure 3. India: Concept 1 inequality calculated in nominal and real terms (1980-81 prices)



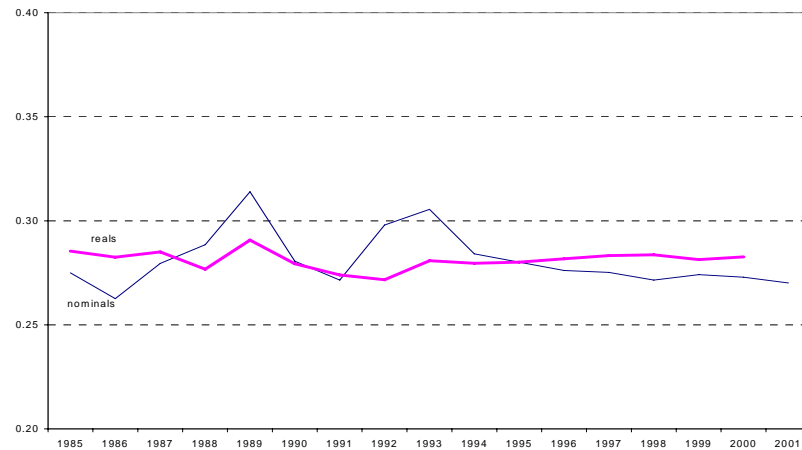
Note: The data on 1999 real GDPs per capita are incomplete and hence not included.

In India, both real- and nominal-based Concept 1 inequality was rising during the last twenty years (Figure 3). However, the same phenomenon of the poorer states' catch-up in terms of price levels is obvious here too as real-based inter-state inequality is higher than the inequality calculated from nominal state GDPs per capita.

In Brazil, inter-state inequality moves up during the periods of high and volatile inflation: between 1986 and 1990 inflation accelerated from 150 percent annually to almost 3000 percent and inter-state inequality, measured by nominal GDPs per capita, rose. (However inequality measured in real terms stayed remarkably constant; Figure 4). Deceleration of inflation in 1991 was accompanied by a decline in inter-state income differences. But during the next three years as inflation accelerated again (from 400 percent in 1991 to more than 2000 in 1994), nominal inter-state inequality rose again. This is an evolution very similar to what we observe for the United States where also inter-state inequality rose during the inflationary period 1977-1981. In both cases, it seems that nominal GDPs per capita are very differently affected from state to state when inflation is high. This is not unexpected. We know that high inflation and hyperinflation are accompanied by greater price volatility, that is real prices of goods and services vary more in high inflation environments (see Parks (1978) and for more recent evidence,

Dabus (2000)). It is then not surprising that different states—which produce different output mixes—will also tend to be affected more unevenly when inflation is high than when it is low.

Figure 4. Brazil: Concept 1 inequality calculated in nominal and real terms (1995 prices)



Note: The state of Tocantins, due to lack of data, is not included in years 1985-1988.

5. Concept 2 regional inequality

In principle population-weighted inequality (Concept 2) is interesting mostly as a stepping stone, or a lower bound to “true” inequality between individuals. As mentioned before, it may reflect the “feeling” of inequality within a country much better than the unweighted Concept 1, particularly when regional inequalities coincide with other types of horizontal cleavages.²⁵ Changes in the population-weighted inequality are important for two additional reasons. First, we may want to see whether inequality is driven by the differing population growth rates between the regions. For example, increasing Concept 2 inequality (and very likely increasing inter-personal inequality) may be due to migration of the population into faster-growing or richer regions.

Secondly, an interesting issue is whether there is correlation between growth rates and population size of different regions. This is in turn relevant for two reasons.

²⁵ As Venables and Kanbur (2003, p. 11) write: “...it may be the case that the distribution of individual attributes such as ability, is the same in all regions. Spatial [Concept 2] inequality is then particularly inequitable and—especially when aligned with political, religious or ethnic tensions—may be dangerous to social and political stability. If spatial divisions align with ethnic, religious or language splits, as so often happens, then the between group component of inequality takes on a greater significance than its contribution to interpersonal (Concept 3) inequality.”

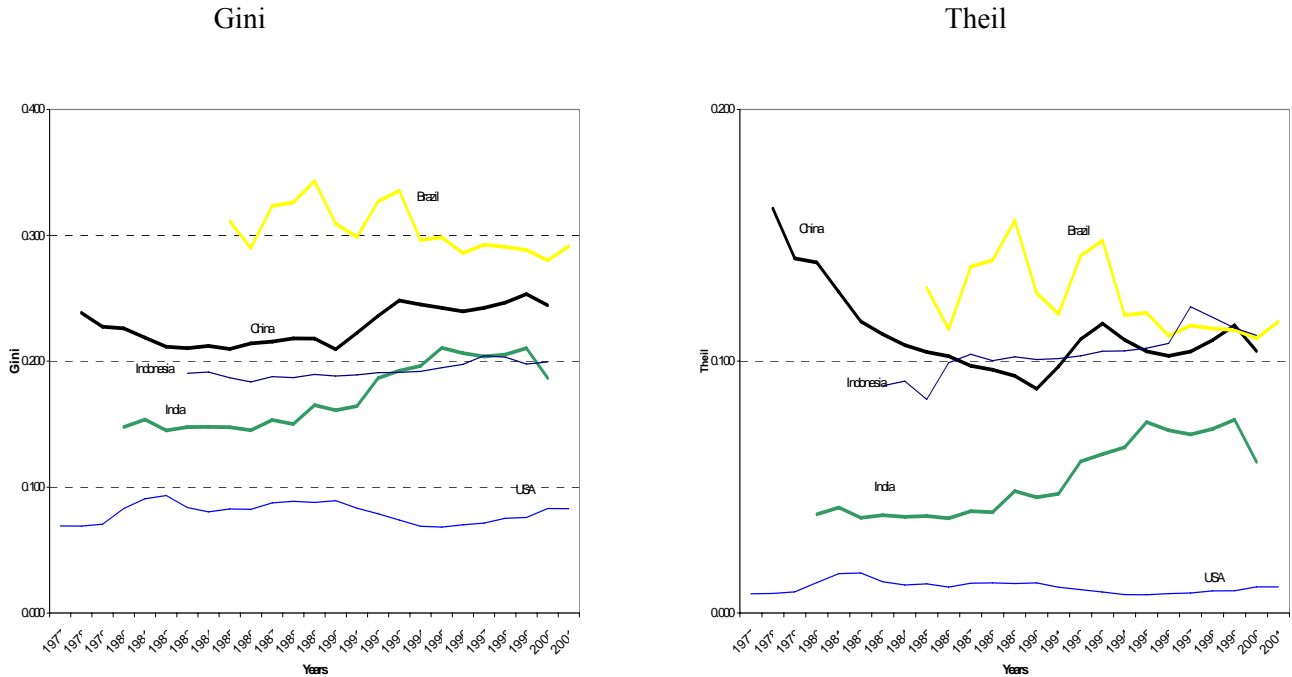
Assuming for the moment that there is no migration between the regions, the evidence of a positive relationship between population and economic growth may help shed additional light on the issue of increasing returns to scale: greater population may help growth if it delays the onset of diminishing marginal returns. Second, higher growth in more populous regions has obvious (positive) implications for the reduction in poverty. If income growth rates tend to be higher in more populous regions then—everything else being the same—poverty reduction will be greater.

We now turn to these two issues: (i) the influence of uneven regional population growth on Concept 2 inequality, and (ii) the association between regional population size and per capita GDP growth. But before we do so, we need to quickly look at the results of Concept 2 inequality for the five countries. They are displayed in Figure 5. Brazil shows the highest inequality with the Gini of about 30. As mentioned before, Concept 2 inequality sets the lower bound to inter-personal (Concept 3) inequality. This means that—were all individual within each state of Brazil to have the same income—overall inequality in Brazil would still be substantial. China has a Gini of about 25, Indonesia and India about 20 (with the latter steadily catching up) and the US less than 10.²⁶ The ranking of the countries is basically unchanged throughout the period. In the 1990's, inequality was on the rise in all three Asian countries. In the 1980's, however, China's inequality—measured by the Gini—is constant while Theil shows a strong downward trend. This indicates that the period must have been characterized by the catch-up of the poorest (population-weighted) provinces.²⁷ For the three Asian countries, there is a clear contrast between the 1980's when regional inequality was declining or constant and of the 1990's when it is also very clearly going up in all three countries.

²⁶ Since size income inequality in the United States is around 40 (see the 2004 release of World Income Inequality Database available at <http://www.wider.unu.edu/>), this means that differences in states' mean incomes explain only about a quarter of total US inequality.

²⁷ This is because Theil is more sensitive than the Gini to what happens at the tails of distribution. A more detailed discussion of China's concept 2 inequality is available in Annex 2 of the Internet version of this paper available at <http://econpapers.hhs.se/RAS/pmi44.htm>

Figure 5. Concept 2 (population-weighted inter-regional inequality) in the five countries



Note : Based on nominal state-level GDP per capita except for Indonesia where only real amounts are available.

We next decompose the change in Gini between the first year for which we have data (late 1970's or early 1980's depending on the country), 1990 and 2000. Table 5 shows what Concept 2 inequality would have been if the state/provincial distribution of population had remained unchanged, and if provincial growth rates had been the same. This enables us to distinguish two effects: uneven population growth, and uneven per capita economic growth. The results show that the effect of differential population movements has been fairly minor in all countries (see columns 4 and 9). Its greatest effect was in the decade of the eighties in Indonesia where it was equalizing and the decade of the nineties in China where, on the contrary, it contributed to greater inequality. One has to be careful however in the interpretation of the population change. It is for example true, in an accounting sense, that had the composition of Chinese population by province remained the same as in 1990, regional inequality in 2000 would have been less. However, this takes the per capita incomes by province as given. Yet in a deeper sense this is wrong. This is because population might have moved in response to higher wages in richer states bringing –let us suppose– in the process the per capita income in the

richer states down. Without such an equilibrating movement of labor, inequalities might have been even greater. Thus for a more meaningful analysis, we would need information on natural vs. mechanical population changes by state/province.

We turn next to the contribution of differential growth rates to the change in Concept 2 inequality (see columns 6 and 11). There the most remarkable turnaround is in the case of China: while in the 1980's, differential growth by province *reduced* regional inequality by almost 3 Gini points, in the next decade, differential growth added 2.6 Gini points to inequality. In the first period, it is the poor and populous provinces that grew relatively fast; in the second period, the poor and populous provinces tended to grow slower than average. This is shown in Figure 6 where we note that the position of many large and relatively poor provinces slipped from being in the NW quadrant (that is, growing faster than average) in the decade of the 1980's, to the SW quadrant (growing slower than average) in the decade of the 1990's.

In India, differential growth was disequalizing throughout although more so in the decade of the nineties. In Indonesia too it was more disequalizing in the 1990's than in the 1980's. It is only in Brazil, in the 1990's, that differential growth was substantially equalizing.

The equalizing effect of differential growth rates can occur for two reasons however. First, poorer regions may grow faster than the rich. Second, although there may be no income convergence among the regions as such, there could be income convergence among the populous regions. This would tend to reduce Concept 2 inequality. To explain: assume that populous regions are uniformly distributed across income distribution and that there is zero correlation between growth rates and initial incomes. Thus, there is neither convergence nor divergence. However, within the subgroup of populous regions, let correlation between initial income and growth may be

negative. These regions' incomes will therefore become more clustered, and since it is them who “matter” for Concept 2 inequality, inequality will go down.²⁸

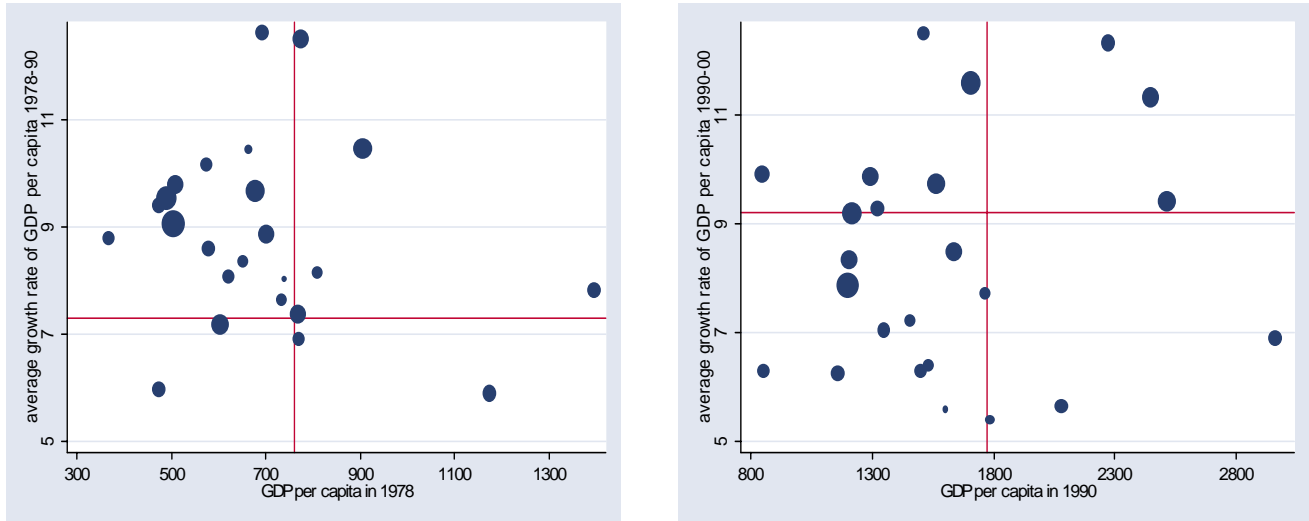
²⁸ Suppose that there are two rich and small regions called A and B, and two poor and populous regions called C and D. Let there be no income convergence, and let A and C grow at high positive rates, and B and D decline. Now, Concept 2 inequality will increase since the two key regions which, because of their sizes determine what happens to Concept 2 inequality, will have become dissimilar. Note that this takes place while there is neither income divergence nor a relationship between population size and growth—but simply income divergence between populous regions.

Table 5. Concept 2 Gini coefficients with actual and hypothetical populations and income

	Change between (circa) 1980 and 1990						Change between 1990 and 2000				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Actual (circa)1980 Concept 2 inequality	Actual 1990 Concept 2 inequality	1990 Concept 2 inequality if population compositio n did not change	Population effect (2)-(3)	1990 Concept 2 inequality if GDP per capita growth rates were the same	Differential growth effect (2)-(5)	Actual 2000 Concept 2 inequality	2000 Concept 2 inequality if population compositio n did not change	Population effect (7)-(8)	2000 Concept 2 inequality if GDP per capita growth rates were the same	Differential growth effect (7)-(10)
USA	6.9	8.9	8.9	0	7.1	+1.8	8.3	8.4	+0.1	8.9	-0.6
China	23.9	20.9	20.9	0	23.8	-2.9	24.4	24.0	+0.4	21.8	+2.6
India	14.8	16.1	16.1	0	14.8	+1.3	18.7	18.7	0	16.2	+2.5
Brazil	31.1	30.9	30.8	+0.1	31.3	-0.4	28.0	28.0	0	31.0	-3.0
Indonesia	19.0	18.8	19.6	-0.8	18.5	+0.3	19.9	20.0	-0.1	18.8	+1.1

Note: The “circa 1980” means the first year for which the data are available: it is 1977 for the United States, 1978 for China, 1980 for India, 1985 for Brazil and 1983 for Indonesia. All Concept2 Ginis calculated from nominal data except for Indonesia. Indonesian data exclude oil and gas component of GDP.

Figure 6. China: Provincial per capita growth, initial income level and population size between 1978 and 1990 (left panel) and 1990 and 2000 (right panel)



Note: Vertical axis shows average growth rate (in percent per capita) over each period. The size of each dot reflects population weight of the province. The city provinces of Beijing, Tianjin and Shanghai are not shown. Horizontal lines drawn at the level of nation-wide average growth during each period (7.3 percent per capita in the first, 9.2 percent per capita in the second). Vertical line drawn at respectively 1978 and 1990 nation-wide GDP per capita (expressed in 1995 international dollars).

To check the second possibility we regress regional rates of growth over a period of five-years on initial regional income levels as in (5)

$$\ln y_{it} - \ln y_{i,t-5} = \beta_0 + \beta_1 y_{i,t-5} + e_{it} \quad (5)$$

where i is i -th region, t is time, y = real GDP per capita, and the sign of the coefficient β_1 indicates presence or absence of unconditional convergence. The regression is population-weighted since our interest is not in convergence as such but in convergence conditional on population size. The regression is calculated for each country and for all years. This is done in order to avoid biasing the results through choice of the time period and stage of the economic cycle.²⁹ The clearest results (see Figure 7) are for India: they show a statistically significant regional divergence since the early 1990's. For China, it is only in the five-year period ending in 1999, that we come close to having a

²⁹ Petrakos et al. (2002, p. 4) for example show how, using the same data, the choice of different time intervals may yield either convergence or divergence.

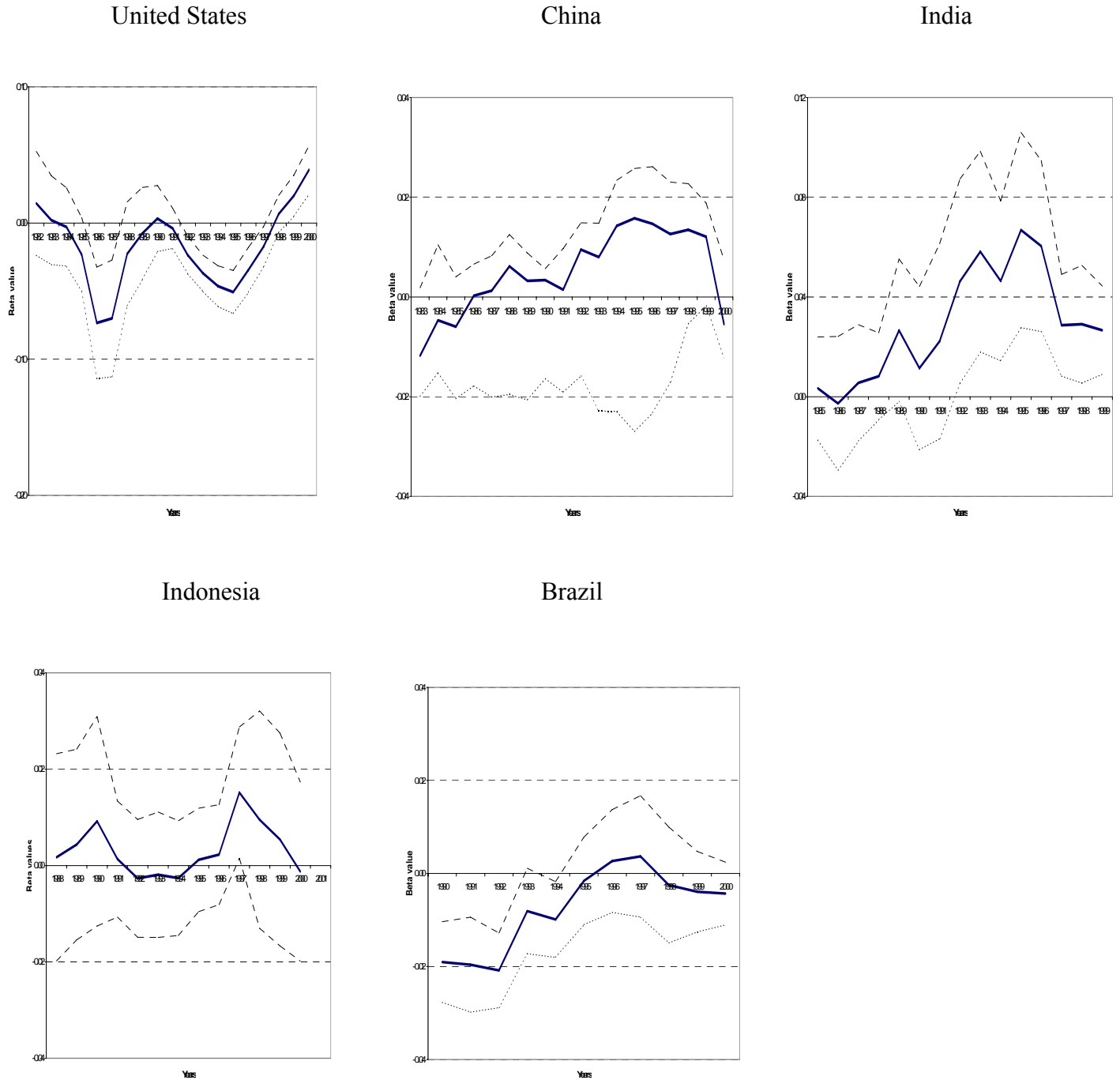
statistically significant regional divergence. Similarly, Indonesia has a five-year period of divergence ending in 1997. In contrast to the three Asian countries that all show some evidence of population-weighted divergence in the 1990's are Brazil and the United States. Brazil's states were strongly converging until the mid-1990's. The United States had a period of convergence from about 1990-91 until the late 1990's.

Consider now what is behind these results. As mentioned before, Concept 2 convergence or divergence is in reality a convergence or divergence within the subset of populous regions. The difference in outcomes among the most populous states is very clear in the case of India. Consider the three most populous Indian states: Uttar Pradesh (population 170 million in 1999), Bihar (107 million) and Maharashtra (95 million). Between 1990 and 1999, GDP per capita in Maharashtra that was the richest of the three states increased by 60 percent (in real terms). Meanwhile, in Bihar—which is the poorest state in India—GDP per capita remained the same. Finally, Uttar Pradesh—which has the median level of income among the states—saw its GDP per capita increase by about 15 percent. The situation is similar in China. The three largest provinces are Sichuan (population 114 million in 2000), Henan (92 million) and Shandong (91 million). During the decade of the nineties, Shandong which was the richest of the three in 1990 saw its real GDP per capita triple. Henan's GDP per capita increased by 2½ times, and Sichuan's (which was the poorest of the three provinces) “only” doubled. In conclusion, in both India and China, the most populous regions diverged in the 1990's and that drove Concept 2 inequality—and possibly inter-personal inequality as well—up.

6. How to explain changes in inter-regional inequality?

As we have seen, levels of inter-regional inequality are quite different in the five countries. Indonesia and Brazil have the highest level of Concept 2 inequality. They are followed by China. India is in the middle but is the only country showing a consistent increase in inter-regional inequality while the United States displays the most uniform levels of GDP per capita between its component states. We shall try to explain within-national convergence or divergence in more general terms, that is by looking at the variables of interest (at national level) that might contribute to or reduce inter-regional inequalities. These variables are not easy to identify in general because of countries' specificities. Thus, for example, the same overall growth rate might produce regional divergence in one case, and regional convergence in another depending on what drives growth. When growth is narrow-based as in case of oil production, the differences between regions (oil-rich and the rest) are likely to increase. When growth is broad-based, or is fuelled by agricultural growth (as in the case of China during the early liberalization), it can be expected to help convergence as poorer and agricultural regions catch up with rich regions. The general presumption—as we have seen from Williamson's (1965) hypothesis—is that growth will tend to be regionally disequalizing (even if ultimately higher income may be associated with lower inequalities). We may thus expect growth rate to be positively related to regional inequality. This is also a finding obtained by Petrakos et al. (2003) in their study of regional inequality in the European Union. The same results are further reported in the case of Indonesia by Akita and Kawamura (2002, p.12) who find that the period of fast growth between 1993 and 1996 was associated with a slightly increasing Concept 2 inequality while the crisis led to a decline in regional inequality (p.16). In a long-run 1939-95 study of Brazil's regional inequality Azzoni (2001, p. 144) finds the same relationship.

Figure 7. Unconditional β regional population-weighted convergence or divergence
(β coefficient with the 95 confidence interval)



Note: the regression is $ROG(t-5,t)=\beta_0+\beta_1 \log GDP(t-5)$ (population weighted) where ROG =annualized per capita growth rate between $t-5$ and t , and GDP = real GDP per capita in year $t-5$ (all expressed in constant international 1995 dollars). Indonesian data exclude oil and gas portion of GDP.

Similarly, the effect of greater openness (approximated by the trade to GDP ratio in nominal terms) may be ambiguous. Openness can help the already rich regions, or can create income gaps where none existed. But it can also help poorer regions whose output (e.g. agricultural goods) was artificially held down through price controls. On the other hand, one can argue that there may be certain policies associated with globalization whose effects are less likely to be ambivalent. They include policies of financial liberalization and higher interest rate that tend to favor rich households. To the extent that rich people are concentrated in certain areas, such policies will increase spatial inequality. As for the empirical results, they span the entire gamut. Kanbur and Zhang (2003) find that openness was associated with rising inequality in China. Zhang and Zhang (2003) similarly decompose Concept 1 inequality between China's states and find that about 20 percent of differences in provincial GDPs per capita can be ascribed to differences in trade shares (p. 57).³⁰ Petrakos et al. (2002, p. 19) however find that openness (defined as regional integration within the European Union³¹) did not have a uniform impact on all EU countries: in some it was associated with greater regional inequality, in others by smaller, and in some had no statistically significant effect at all. Wei and Wu (2001) using urban/rural ratios of mean income for more than 100 cities and their adjoining areas in China find that in the period 1988-93 increased openness tended to reduce the urban/rural ratio. Since urban/rural differences are perhaps the principal explanation for regional inequality in China, increased openness would seem to reduce regional inequality.

Visual inspection of Figure 8 which displays the relationship between openness and Concept 1 inequality shows that all of our five countries (with the exception of Indonesia), have, as we would expect, relatively low openness. The graph also shows that in China, increased openness was not associated with increased regional inequality while, on the contrary, in India, this seems to be clearly the case. At the other extreme is

³⁰ Their approach is interesting because they run production functions of the same form for all provinces. One of the arguments in the production function is trade ratio (in addition to education, domestic and foreign capital accumulation etc.) Then log variance of GDPs per capita is decomposed and the regression coefficients times the covariance between income and each argument gives an estimate of that particular argument's contribution to inter-provincial income inequality.

³¹ And measured as share of European trade in country's total trade.

Indonesia where increased openness was associated with a decrease in regional inequality. Finally in Brazil and the United States, there was no apparent relationship between openness and regional inequality. The conclusion is that the country experiences differ and that openness as such may not have the same discernable effects on countries regardless of their level of development, type of economic institutions, and other macro economic policies.

A somewhat different picture however emerges when we look at the relationship between Concept 2 inequality and openness (Figure 9). It is now very clear that in both China and India, increased openness was associated with greater regional inequality. In Indonesia, the relationship is the opposite and in Brazil and the US the correlations seems non-existent.

We then run two types of regressions across both Concept 1 and Concept 2 inequalities (see equation 6). The first type are fixed-effect panel regressions, the second are Arellano-Bover GMM (Generalized Method of Moments) regressions.

$$I_{it} = \beta_0 + \beta_1 ROG_{it} + \beta_2 \ln GDP_{it} + \beta_3 OPEN_{it} + \beta_4 RINT_{it} + \beta_5 INFL_{it} + u_{it} \quad (6)$$

where I = inequality concept that can be either Gini or Theil (with e.g. Gini2 denoting Gini coefficient under Concept 2 inequality), i denotes country, t = year, ROG = real annual rate of growth of per capita GDP, GDP = level of GDP per capita, $OPEN$ = trade to GDP ratio in current prices, $RINT$ = real interest rate (on deposits) in percent per annum, $INFL$ = the average annual rate of inflation proxied by the change in consumer price index. GDP and rate of growth of GDP are included, as discussed in Section 2, to test for the presence of short- and long-term effects of growth on regional inequality (as implied by, among others, by Williamson's 1965 hypothesis).³² The possible role of openness is explained above. We also include real rate of interest and inflation as possible controls on the assumption that both can contribute to regional

³² This is also the reason why a squared GDP per capita term (in the spirit of the Kuznets hypothesis) is not included. The latter is a test of inter-personal inequality in the presence of structural change (and increase in income). Here we test regional inequality which, according to the Williamson and Perroux hypotheses, should be less in the already developed economies where market integration is more advanced. Per capita income thus serves as a proxy for the extent of integration of capital, labor and goods markets.

inequality: the first because a higher rate of interest will increase incomes of property-owners who also tend to be spatially concentrated; the second, because higher inflation is often associated with greater variability in relative prices. To the extent that different regions specialize in producing different output-mixes, then relative price variability will affect them differently, thus adding to inequality.

All variables in (6) are defined at national level. Inequality that is calculated across provinces is Concept 1 or Concept 2 inequality at the *national* level (in the same way that inequality across individuals in a country is a national level variable). The other variables are obviously so. It is an unbalanced panel since the number of observations for each country is not the same (e.g. for China we have the data for the 1978-2000 period, but for Indonesia only for 1983 to 2001). Since there are unobservable country-specific effects and the number of regional units per country varies (which in turn means that, everything else being the same, countries with a greater number of regional units would tend to have higher inequality), we use the fixed effects model. However if we acknowledge that inequality is a process characterized with a high degree of persistence, we have to move to dynamic panels (with the lagged inequality value on the RHS). The introduction of the lagged dependent variable makes the fixed-effect estimator inconsistent in small samples. To account for that we use Arellano-Bover GMM system estimator which similarly to the Arellano-Bond estimator involves first-differencing to remove possible country fixed effects, instrumenting of predetermined and endogenous variables by their lagged values,³³ but also extracts as much information as possible from the data by estimating the regressions both in levels and changes and imposing the equality of the slopes between the two.

³³ We also use as instruments population size and land area. The Sargan test of identifying restrictions is statistically insignificant throughout implying acceptance of exogeneity of instruments.

Figure 8. The relationship between openness and Concept 1 inequality

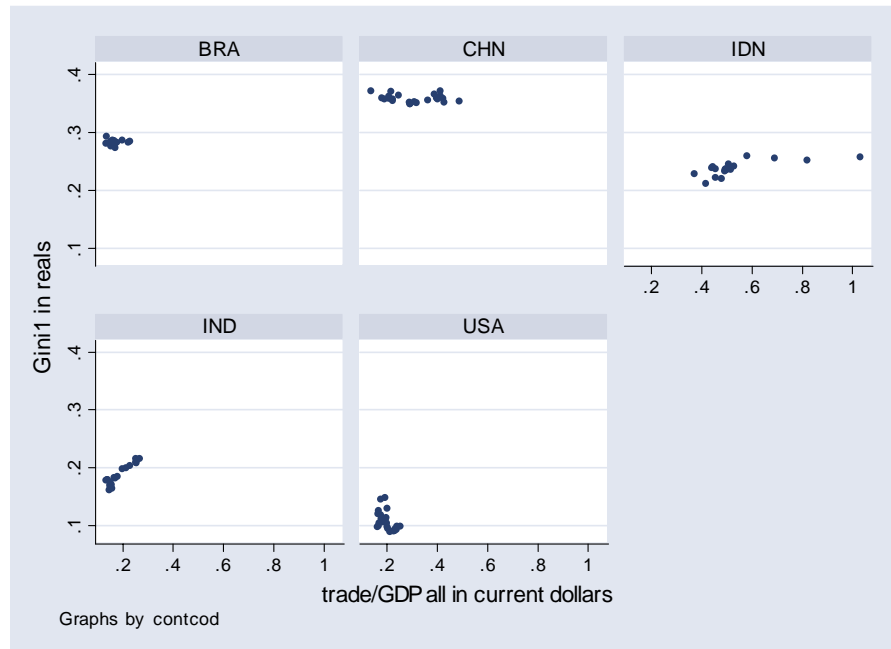
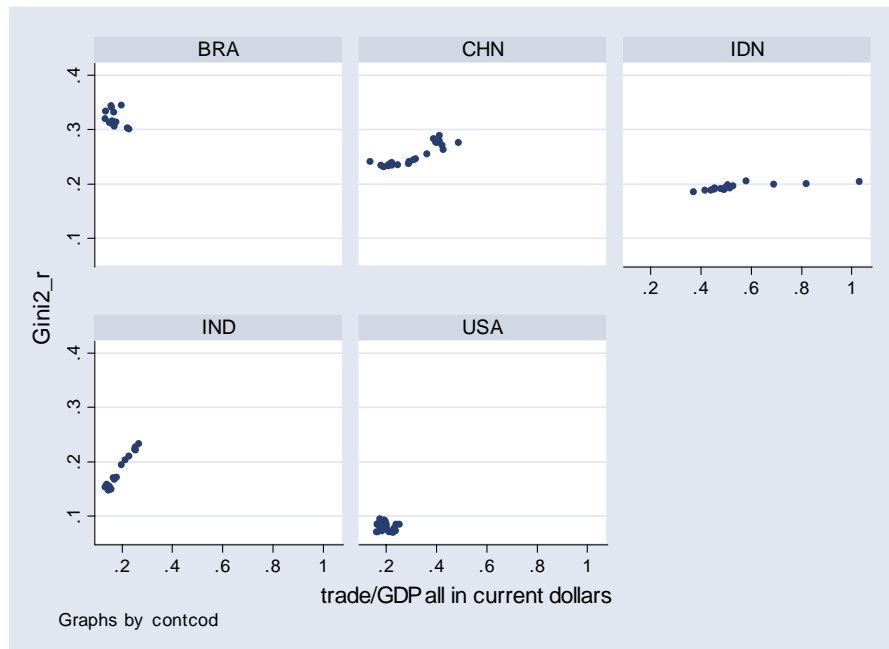


Figure 9. The relationship between openness and Concept 2 inequality



Note: Both inequality concepts are expressed in real terms. The data for Indonesia exclude gas and oil. CHN=China; IND=India, IDN=Indonesia; BRA=Brazil.

Table 6. Determinants of regional inequality

	Concept 1 inequality				Concept 2 inequality			
	Gini1		Theil1		Gini2		Theil2	
	Fixed effects	Arellano-Bond	Fixed effects	Arellano-Bond	Fixed effects	Arellano-Bond	Fixed effects	Arellano-Bond
Lagged Gini or Theil		0.977 (0.000)		0.973 (0.000)		0.973 (0.000)		0.962 (0.000)
ROG	-0.020 (0.69)	0.031 (0.19)	0.0014 (0.98)	0.036 (0.15)	0.028 (0.61)	0.046 (0.006)	0.038 (0.32)	0.004 (0.002)
GDP per capita	0.009 (0.37)	-0.0014 (0.15)	-0.003 (0.77)	-0.001 (0.22)	0.056 (0.000)	-0.002 (0.011)	0.03 (0.000)	-0.001 (0.006)
Openness	0.039 (0.13)	0.004 (0.50)	0.034 (0.15)	0.003 (0.57)	0.005 (0.87)	0.000 (0.99)	0.021 (0.27)	0.003 (0.39)
Inflation	0.0007 (0.60)	0.0002 (0.75)	0.0008 (0.52)	-0.0001 (0.80)	0.004 (0.016)	0.0004 (0.44)	0.003 (0.01)	0.0001 (0.94)
Real interest	0.009 (0.85)	0.014 (0.58)	0.01 (0.81)	0.22 (0.39)	-0.028 (0.57)	-0.01 (0.56)	-0.02 (0.57)	0.004 (0.78)
Constant	0.126 (0.14)	0.014 (0.15)	0.116 (0.14)	0.01 (0.24)	-0.297 (0.002)	0.002 (0.011)	-0.19 (0.003)	0.013 (0.008)
R ² within	0.14		0.07		0.41		0.39	
F	2.5 (0.04)	2010 (0.000)	1.1 (0.38)	1534 (0.00)	10.0 (0.000)	3799 (0.000)	9.4 (0.00)	2540 (0.000)
No. of observations	84	78	84	78	84	78	84	78
Sargan test		74.8 (0.95)		75.4 (0.94)		73.7 (0.95)		78.6 (0.86)

Note: *p* values between brackets. Coefficients significant at 5 percent or less are shaded. Real interest = real annual rate of interest in percent p.a. divided by 1000. Inflation = $\ln(1 + \text{annual inflation rate in percent})$. Openness = trade/GDP (in nominal US dollars). GDP per capita = \ln GDP per capita in 1995 international dollars. ROG = annual rate of growth of GDP per capita (expressed in fractions, e.g. 3 percent is expressed as 0.03). Gini and Theil are expressed as ratios and calculated across real regional incomes. Arellano-Bond is one-step GMM estimator; calculations done using David Roodman's xtabond2.ado STATA software.

The results of estimation are displayed in Table 6. Regarding Concept 1 inequality, we see that practically no variable in either formulation seems to have any influence and the quality of the overall results is very weak. Contrary to several other studies, we do not find evidence that higher rate of growth has a disequalizing effect on regional incomes. The situation is different when we move to the determinants of Concept 2 regional inequality. There, according to both Gini and Theil, and when using dynamic panel specification, higher rate of growth is disequalizing, and higher income level is equalizing. We thus find the same results as obtained by Petrakos, Rodriguez-Pose and Roviolis (2003), in the context of the European Union regions.³⁴ The effect of higher growth rate on Concept 2 inequality is rather small though: an acceleration of national growth rate by 1 percentage point is associated on average with an increase of Concept 2 inequality by 0.046 Gini points. Inflation is disequalizing only in the static (fixed-effects) framework. Openness, however, although it enters all the regressions with a positive sign, is not statistically significant in none of them. The same is true for real rate of interest which shows a sign change as well.

7. Conclusions

The three Asian countries show much greater evidence of increasing regional income inequality than Brazil and the United States. Divergence in China, India and Indonesia is evident in the decade of the 1990's. The clearest case is that of India. In the 1990's, it shows a consistently increasing regional inequality of both Concept 1 and 2 types—whether it is measured by the Gini or by the Theil index. The largest states (Maharashtra, Uttar Pradesh and Bihar) are diverging in their incomes, and generally richer states have registered faster growth.

The experience of the 1990's is in contrast to that of the previous decade for all three countries. In China during the 1980's, there was a convergence between the regions. It coincided with the period of agricultural reforms which enabled poorer regions to catch

³⁴ In fixed-effects though, income level is positively associated with Concept 2 inequality and rate of growth is not significant.

up. Moreover, over the same period, there was also a strong Concept 2 convergence. The process comes to a halt in the 1990's with provincial mean incomes growing at relatively similar rates (Concept 1 inequality stable), but, importantly, Concept 2 inequality increasing. Similarly to what we noted for India, in China the divergence in the second half of the 1990's was driven by the divergence in outcomes between populous provinces. As there is a contrast between Maharashtra and Bihar in India, there is a contrast between Shandong and Sichuan in China.

As for the two non-Asian countries studied here, the United States, despite already being most regionally homogeneous, has generally displayed tendency toward income convergence according to both inequality concepts. Brazil shows least consistent change. There is only a slight change in both Concept 1 and Concept 2 inequality, but it is overshadowed by a cyclical effect of inflation. It is therefore difficult to see whether there is any trend. Nevertheless Brazil has the highest Concept 2 Gini inequality of all countries included here. To see how large that inequality is, one need simply realize that were all incomes within each state absolutely equal, Brazil's overall Gini would still be around 30, a moderate nation-wide inequality level.

When we try to fit our findings into an admittedly Procrustean bed of regression analysis, we do not find much evidence, even after adjusting for countries' unobservable characteristics, that the most obvious variables used in the work on regional inequality explain its Concept 1 change over time. Turning to population-weighted regional inequality, we find that it tends to be positively associated with faster growth and negatively with income level. As for openness, although its coefficient is positive, it is not statistically significant in any of the formulations. High inflation is (in static panel framework) associated with higher Concept 2 inequality.

Global inequality changes and inequality changes in the five most populous countries in the world during the last two decades have indeed been quite dramatic but also extremely complex. Note the following facts. While at the level of these countries, we find no evidence for the narrowing of (unweighted) inter-regional income differences

during the decade of the 1990's, Concept 2 inequality however tended to rise in the two most important countries (India and China) as well as Indonesia. But at the global level, a process directly opposite to the one at the national level was simultaneously taking place. It consisted of widening income differences between the countries (as poor countries have tended to lag in growth rates behind the rich) and a reduction in Concept 2 inequality thanks mostly to China's phenomenal growth. In other words, while within India or China during the last ten years or so, Concept 1 inequality was stable and Concept 2 inequality was increasing, on the global level (mostly thanks to what was happening precisely in these two countries), Concept 1 inequality was up and Concept 2 inequality went down.

Even without addressing the most difficult task of measurement of inter-personal inequality at the global level, one can easily see how contradictory are recent changes in different facets of inequality and how inadequate are simple answers regarding both the direction of change and even more so regarding the causality. In other words, when we ask a question as apparently simple as "what was a change in inequality in country X", we need to define very carefully what we mean by inequality since its different versions may easily move in the opposite directions.

REFERENCES

- Akita, Takahito and Lukman, Rizal Affandi (1995), "Interregional inequalities in Indonesia: A sectoral decomposition analysis for 1975-92", *Bulletin of Indonesian Economic Studies*, vol. 31, No. 2, August, pp. 61-81.
- Akita, Takahito and Kazumi Kawamura (2002), "Regional income inequality in China and Indonesia: A comparative analysis", Paper presented at the 46th Congress of European Regional Science Association, Dortmund, Germany, August 27-31, 2002. Available at <http://www.ersa.org/> and <http://econpapers.hhs.se/paper/wiwwiwsa/ersa02p432.htm>.
- Akita, Takahito and Mitsuhiro Kataoka (2003), "Regional income inequality in the post-war Japan", paper presented at the 47th Congress of European Regional Science Association, Jyväskylä, Finland, August 27-30, 2003. Available at <http://econpapers.hhs.se/paper/wiwwiwsa/ersa03p480.htm>.
- Azzoni, Carlos R. (2001), "Economic growth and regional income inequality in Brazil", *The Annals of Regional Science*, vol. 35, pp. 133-152.
- Barro, Robert and Xavier Sala-i-Martin (1992), "Convergence", *Journal of Political Economy*, vol. 100 (2), pp. 223-251.
- Bliss, Christopher (1999), "Galton's Fallacy and Economic Convergence", *Oxford Economic Papers*, vol. 51, pp. 4-14.
- Bhalla, Ajit S., Shujie Yao and Zongyi Zhang (2003), "Causes on inequalities in China", *Journal of International Development*, vol. 15, pp. 939-955.
- Chen, Shaohua and Yan Weng (2001), "China's growth and poverty reduction: Recent trends between 1990 and 1999", World Bank Research Working Paper No. 2651, July.
- Dabus, Carlos (2000), "Inflationary Regimes and Relative Price Variability: Evidence from Argentina" *Journal of Development Economics* vol. 62, No. 2 (August 2000): pp. 535-47
- Dreze, Jean and P.V. Srinivasan (2000), "Poverty and inequality in India: Evidence from regional data", *Journal of Quantitative Economics*, vol. 16, No. 1, January, pp. 33-50.
- Easterly, William and Ross Levine (2001), "It's Not Factor Accumulation: Stylized Facts and Growth Models", *World Bank Economic Review*, vol. 15(2), pp. 177-220.
- Goerlich, Francisco J. and Matilde Mas (2001), "Inequality in Spain, 1973-91: Contribution to a Regional Database", *Review of Income and Wealth*, Series 47, No.3, September.
- Heston, Alan (2001), "Treatment of China in PWT 6.1", December. Available <http://pwt.econ.upenn.edu/>.
- Hirschman, Albert (1958), *The strategy of economic development*, New Haven: Yale University Press.
- Holz, Carsten A. (2004), "China's Statistical System in Transition: Challenges, Data Problems and Institutional Innovations", *Review of Income and Wealth*, Number 3, September, pp. 381-409.
- Jha, Raghendra (2004), "Reducing poverty and inequality in India: has liberalization helped", in Giovanni Andrea Cornia (ed.) *Inequality, growth and poverty in an era of liberalization*, Oxford, U.K. and New York: Oxford University Press. Also available at <http://rspas.anu.edu.au/economics/publish/papers/wp2002/wp-econ-2002-04.pdf>.
- Kanbur, Ravi and Xiaobo Zhang (2003) "Fifty years of regional inequality in China: A journey through central planning, reform and openness". Available at <http://people.cornell.edu/pages/sk145/papers/Halfcentury81.pdf>. forthcoming in *Review of Development Economics*, vol. 9, No. 1, 2005.
- Krugman, Paul (1991), *Geography and Trade*, London: MIT Press.
- Krugman, Paul and Anthony Venables (1995), "Globalization and the inequality of nations", *Quarterly Journal of Economics*, vol. 110, No. 4, pp. 857-880.
- Maddison, Angus (1998), *Chinese Economic Performance in the Long Run*, Paris: OECD Development Centre.
- Maddison, Angus (2003), *The World Economy: Historical Statistics*, Paris: OECD Development Centre Studies.
- Maddison, Angus (2003), *The World Economy: Historical Statistics*, Paris: OECD Development Centre Studies.
- Milanovic, Branko (2002), "True world income distribution, 1988 and 1993: First calculation based on household surveys alone", *The Economic Journal*, vol. 112, No. 476, January 2002, pp. 51-92.

- Milanovic, Branko (2005), *Worlds Apart: Global and International Inequality 1950-2000*, Princeton, N.J.: Princeton University Press.
- Myrdal, Gunnar (1957), *Economic theory and underdeveloped regions*. London: Hutchinson.
- Perroux, Francois (1970), "A note on the concept of growth pole" in D. McKee, R. Dean and W. Leahy (eds.), *Regional economics theory and practice*, New York: Free Press, pp. 93-103.
- Perroux, Francois (1988), "The Pole of Development's New Place in a General Theory of Economic Activity", *Regional economic development: Essays in honour of Francois Perroux*, Boston; London; Sydney and Wellington, New Zealand: Unwin Hyman, pp. 48-76
- Petrakos, George, A. Rodriguez-Pose and A. Rovolis (2003), "Growth, integration and regional inequality in Europe", paper presented at the 47th Congress of European Regional Science Association, Jyväskylä, Finland, august 27-30, 2003.
Available at <http://www.ersa.org/ersaconfs/ersa03/cdrom/papers/46.pdf>.
- Parks, Richard (1978), "Inflation and relative price variability", *Journal of Political Economy*, vol. 86, pp. 79-95.
- Ram, Rati (1992), "Interstate income inequality in the United States: measurement, modelling and some characteristics", *Review of Income and Wealth*, vol. 38, pp. 39-49.
- Rawski, Thomas (2001). "What's Happening to China's GDPs Statistics", *China Economic Review*, No. 4, vol. 12., pp. 347-354.
- Renard, Marie-Francoise (2002), "A Pessimistic View on the Impact of Regional Inequalities", *China Economic Review* vol. 13, No.4, pp. 341-44
- Rosenstein-Rodan, Paul (1943), "Problems of industrialisation of Eastern and Southeastern Europe", *Economic Journal*, vol. 53, pp. 202-211.
- Sharping, Thomas (2001), "Hide-and-seek: China's elusive population data", *China Economic Review*, No. 4, vol. 12., pp. 323-332.
- Shorrocks, Anthony and Guaghua Wan (2004), "Spatial decomposition of inequality", WIDER Discussion Paper No. 2004/01. Available at <http://www.wider.unu.edu/publications/publications.htm>.
- Venables, Anthony J. and Ravi Kanbur (2003), "Spatial Inequality and Development", Department of Applied Economics and management working paper 2003-38, Cornell University, November. (Republished as introduction to Venables and Kanbur (eds), *Spatial Development and Inequality*, Oxford: Oxford University Press, 2005).
- Wei, Shang-jin and Yi Wu (2001). "Globalization and Inequality: Evidence from Within China." *Center for Economic Policy Research Discussion Paper*, December, No. 3088.
- Williamson, Jeffrey (1965), "Regional inequality and the process of national development: a description of patterns", *Economic development and cultural change*, vol. 13. No. 4, pp. 3-45.
- Xu, Xianchun (1999), "Evaluation and Adjustments of China's Official GDP by the World Bank and Prof. Maddison", *Journal of Econometric Study of Northeast Asia*, vol. 1, No.2.
- Yao, Shujie and Zongyi Zhang (2001), "On regional inequality and diverging clubs: A case study of contemporary China", *Journal of Comparative Economics*, vol. 29, pp. 466-484.
- Yemtsov, Ruslan (2002), "Quo Vadis: Inequality and Poverty Dynamics Across Russian Regions", World Bank, mimeo.
- Zhang, Zongyi, Aying Liu and Shujie Yao (2001), "Convergence of China's regional incomes 1952-1997", *China Economic Review*, vol. 12, pp. 243-258.
- Zhang, Xiaobo and Kevin H. Zhang (2003), "How does globalisation affect regional inequality within a developing country? Evidence from China", *Journal of Development Studies*, vol. 30, No.4, April, pp. 47-67.

Annex 1. Administrative organization of the five countries and data coverage

China

China is officially administratively divided into 34 regional units: 23 provinces (including Taiwan, China), 2 special administrative regions (Hong Kong and Macau), 5 autonomous regions (Guangxi, Inner Mongolia, Ningxia, Xinjiang and Tibet) and 4 municipalities (Beijing, Shanghai, Tianjin and Chongqing). The autonomous regions differ from the rest because of the presence of significant non-Han minorities in them.

We are not including in the analysis Taiwan (China), Hong Kong and Macau because the GDP data for them are not shown together with other regions. In addition, Chongqing acquired the status of a separate unit (municipality) only in 1997 and in our data is included together with Sichuan. This therefore gives us 30 regional units which for simplicity we call “provinces.” Out of these thirty, full data sets (both nominal and real GDPs) are available for 27 provinces which represent 99 percent of Chinese population (not counting Taiwan, China; Hong Kong; Macau). For two provinces (Qinghai and Hainan), nominal GDP is available from 1985 onwards.³⁵ Hence in calculations with nominal values, 29 provinces are included. Finally, for Tibet neither real nor nominal GDP data are available.

A note is in order regarding Chinese GDP per capita data.³⁶ There are three types of problems with these data. First, all-China GDP data are the subject of a long academic dispute. Many authors argue that Chinese official statistics exaggerate both the level and rate of growth of the economy (Maddison, 2003, p. 151; Maddison, 1998; Heston, 2001; Xu 1999; Rawski, 2001). If this is true then provincial growth rates must also be exaggerated. While we tend to believe that these authors do have a point and that the total GDP correction does make sense, the only source of *provincial* Chinese statistics is the official State Statistics Bureau. We thus use official regional and nation-wide statistics

³⁵ Hainan became a separate province in 1988. Previously it was part of Guangdong province.

for the entire 1978-2001 period. Second, until 1994, the sum of provincial GDPs was approximately equal to the official value of the nation-wide GDP. The discrepancy was within a very narrow range of 1-2 percentage points possibly due to the mistakes of classification. But after 1994, the sum of provincial GDPs is systematically greater (up to 15 percent) than the nation-wide GDP (see Holz, 2004, pp. 388-9; Milanovic, 2005, pp. 97-8). It is unclear what the source of this large discrepancy is. Heston (2001, p. 3) ascribes it to the “winds of falsification”. Holz (2004) similarly blames data falsification at lower-level administrative units. Third, the denominator in GDP per capita data is the officially registered population (see Scharping, 2001). This tends to underestimate actual population living in richer provinces which, due to their dynamism, attract the bulk of “floating population” (people who do not have official permits to live in a given city). GDP per capita for richer provinces is accordingly overestimated and GDP per capita for poorer provinces underestimated, thus biasing our inequality statistics upward. We have no way out of this conundrum. This is because the only existing data on provincial GDP and provincial population are the official data.

India

India was until the year 2000 administratively divided into 25 states and 7 Union territories. The Union territories are very small with the exception of the federal capital of Delhi. In total they account for about 11 million people or a little over 1 percent of total population (in the 1991 census). India is officially defined as a “union of states” (Article 2 of the 1949 Constitution). Out of 25 states, we have GDP per capita data for 14 states. These are the largest states comprising 92 percent of total Indian population in 2000. Of the states for which we do not have the data, the largest are Assam, and Jammu and Kashmir (respectively 23 and 8 million according to the 1991 census). The disputed parts of Kashmir are also omitted from the analysis (the data for them are included in Chinese and Pakistani statistics).

³⁶ An entire issue of *China Economic Review* (vol. 12, No. 4, 2001) was dedicated to the data problems in Chinese statistics.

In 2000, three new states were created. They are Uttaranchal or Uttarkhand (population of 4 million), carved out of Uttar Pradesh (170 million before the split), Jharkhand (population of 15 million) carved out of Bihar (107 million before the split), and Chhattisgarh (population of 19 million) carved out of Madhya Pradesh (79 million before the split). All three are based on ethnic distinctiveness from their “mother” states. In our data for 2000, these states are included as part of their “mother” states.

United States

United States is administratively divided in 50 states and 1 federal district (the federal capital of Washington). In the Constitution, the United States is not officially defined as a federation although this is implicit from the text. United States also controls some 12 Commonwealths and territories out of which Puerto Rico is the largest (3.7 million people). Our data cover the 50 states and the District of Columbia. However, the latter, due to its peculiarities,³⁷ is not included in the analysis which thus bears only on the 50 federal units. Since District’s population is only about 600,000 people, our coverage of the United States is almost complete.

Brazil

Brazil is administratively divided in 26 states and 1 federal district (the federal capital of Brasilia). It is officially a federation (hence the official name since 1988 of “Federative Republic of Brazil”). Our data include all states with the exception of Tocantins for the years 1985-88. The federal district of Brasilia (population 2 million) is not included. Thus the population coverage is about 99 percent.³⁸

Indonesia

³⁷ For example, District’s GDP per capita is unrealistically high, at more than \$PPP 90,000. This is due to the fact that many businesses are registered in the District while people who work or own them live just outside the borders of the District.

³⁸ The state level value added data for Brazil (available at Instituto Brasileiro de Geografia e Estatística, <http://www.ibge.gov.br/english/estatistica/economia/contasregionais/default.shtm>) account for between 95 and 98 percent of all-Brazil sum from the official statistics. Almost all of that difference is explained by the fact that state data used here exclude the federal territory of Brasilia (its share of all-Brazil GDP varies between 3 and 5 percent).

Indonesia is administratively divided in 26 provinces. Before East Timor's independence in 1999, there were 27 provinces. The data for East Timor are available for the entire period during which it was an Indonesian province, but for the reasons of comparability, we have omitted it from our calculations throughout the entire period. The population coverage of Indonesia is thus 100 percent.

Included countries and states/provinces

China	India	Indonesia	Brazil	USA
(Prices: 1978)	(Prices: 1980-81)	(Prices: 1983)	(Prices: 1985)	(Prices: 1995)
Beijing	Andhra Pradesh	Aceh	Acre	Alabama
Tianjin	Bihar	North Sumatra	Alagoas	Alaska
Heibei	Gujarat	West Sumatra	Amapa	Arizona
Shanxi	Haryana	Riau	Amazonas	Arkansas
Neimeng	Karnataka	Jambi	Bahia	California
Liaoning	Kerala	South Sumatra	Ceara	Colorado
Jilin	Madhya Pradesh	Bengkulu	Distrito Federal	Connecticut
Heilongjiang	Maharashtra	Lampung	Espirito Santo	Delaware
Shanghai	Orissa	Jakarta	Goias	Florida
Jiangsu	Punjab	West Java	Maranhao	Georgia
Zhejiang	Rajasthan	Central Java	Mato Grasso	Hawaii
Anhui	Tamil Nadu	Yogyakarta	Mato Grosso do Sul	Idaho
Fujian	Uttar Pradesh	East Java	Minas Gerais	Illinois
Jiangxi	West Bengal	Bali	Para	Indiana
Shandong		West Kalimantan	Paraiba	Iowa
Henan		Central Kalimantan	Parana	Kansas
Hubei		South Kalimantan	Pernambuco	Kentucky
Hunan		East Kalimantan	Piaui	Louisiana
Guangdong		North Sulawesi	Rio de Janeiro	Maine
Guangxi		Central Sulawesi	Rio Grande do Norte	Maryland
Hainan		South Sulawesi	Rio Grande do Sul	Massachusetts
Sichuan		Southeast Sulawesi	Rondonia	Michigan
Guizhou		West Nusa Tenggara	Roraima	Minnesota
Yunnan		East Nusa Tenggara	Santa Catarina	Mississippi
Shannxi		Maluku	Sao Paulo	Missouri
Gansu		Irian Jaya	Sergipe	Montana
Qinghai		Timor Timur	Tocantins	Nebraska
Ningxia				Nevada
Xinjiang				New Hampshire
				New Jersey
				New Mexico
				New York
				North Carolina
				North Dakota
				Ohio
				Oklahoma
				Oregon
				Pennsylvania
				Rhode Island
				South Carolina
				South Dakota
				Tennessee
				Texas
				Utah
				Vermont
				Virginia
				Washington
				West Virginia
				Wisconsin
				Wyoming

Annex 2. Regional inequality in China

China at the end of the 20th century: different regional inequality break-downs
(Concept 2 inequality throughout)

	Akira and Kawamura (2002)	Kanbur and Zhang (2003, Table 2)	Modified (Milanovic) Kanbur and Zhang (2003)	Bhalla, Yao and Zhang (2003)	Milanovic (here)
Year	1998	2000	2000	1995	2000
Welfare concept	GDP per capita	Mean real per capita consumption 1/	Mean real per capita consumption 1/	Mean real per capita consumption 2/	GDP per capita
Smallest unit with data	District level GDP per capita	Rural/urban	Rural/urban	Rural/urban (=peasant/non-peasant)	Provincial GDP per capita
Aggregation	By province (26) and by regions (3)	By province (28)	By province (28)	By province (28)	---
Total number of observations (per year)	335	56	56	56	29
Concept 2 Theil 3/ (which Theil?)	24.9 Theil (1)	24.8 Theil (0)	24.8 Theil(0)	16.54 (see note) Theil (1)	10.4 Theil (1)
Between 3 large regions	6.6 (27)				
Between provinces (within each region)	2.8 (11)				
Between 28 provinces alone	9.4 (38)		10.9 (44)	11.0 (67)	10.4 (100)
Within provinces (between districts)	15.6 (62)				
Within provinces (between rural and urban)			13.9 (56)		
Between overall rural and overall urban mean		13.9 (56)		10.4 (75)	
Within rural areas and within urban areas		10.9 (44)		6.1 (25)	
Within rural areas				3.8 (21)	
Within urban areas				2.3 (4)	
Concept 2 Gini 3/	n.a.	37.2	37.2	n.a.	24.4

Sources: Takahiro and Kawamura (2002, Table 1, p. 26). Kanbur and Zhang (2003, Table 2, p. 27). Bhalla, Yao and Zhang (2003, Table 2, p. 947).

Modified Kanbur Zhang consists of taking Kanbur and Zhang data and applying somewhat different partitions (for explanation, see the text below).

1/ In 1983 prices. 2/ In 1990 prices, using provincial price deflators. 3/ Based on the most detailed partition used in the paper, e.g. 335 observations in Akira and Kawamura paper etc.

Note: Theil (1) is Theil entropy index. Theil (0) is mean log deviation index. The definitions are:

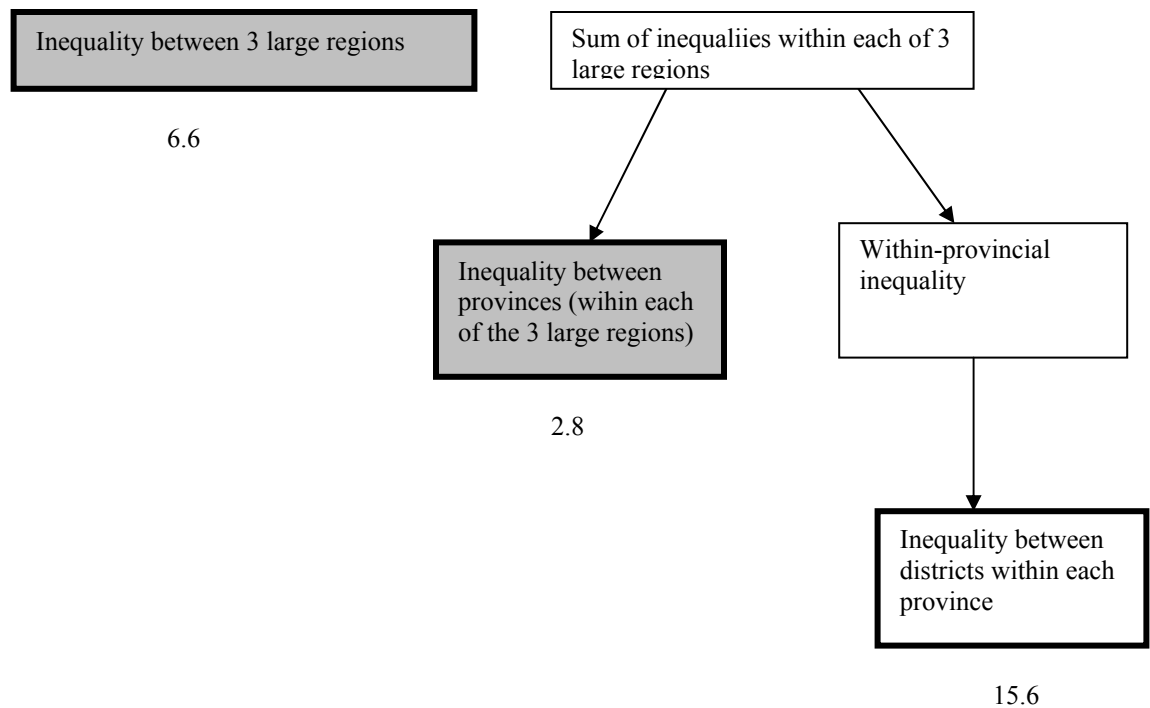
$$Theil_1 = \sum_{i=1}^n p_i \frac{y_i}{\mu} \log \frac{y_i}{\mu}$$

$$Theil_0 = \sum_{i=1}^n p_i \log \frac{\mu}{y_i}$$

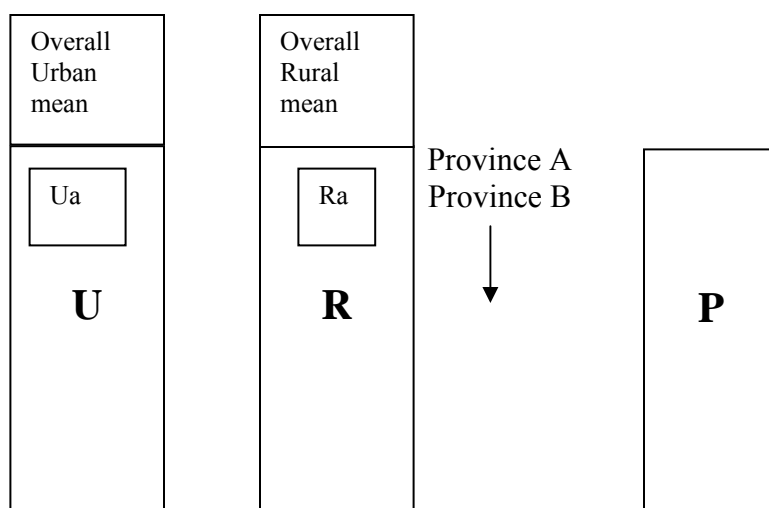
Note: The Kanbur and Zhang (2003) and Bhalla, Yao and Zhang (2003) results should be very similar since their welfare concepts and the aggregations/partitions are practically the same. And indeed, for the year 1995 (last year in Bhalla et al.) Kanbur and Zhang report a Concept 2 Theil of 17.7 while Bhalla et. al. have obtain a Theil of 16.5 (see the results in the Table here). However, inexplicably, in their paper Bhalla et. al (Table 2, p. 947) show the total Concept 2 Theil of 13.87 while the sum of the three element into which the Theil is decomposed yields 16.54. I think that the latter amount is correct.

The decomposition rules can be also presented graphically. In the Akita and Kawamura (2002) paper, total inequality is the sum of three inequalities given in “bolded” boxes with their respective values in the year 2000 given underneath the boxes. We can now easily see that provinces *within* each of the three large regions are fairly homogeneous (in terms of per capita income) and that the bulk of inequality is concentrated at the provincial level (differences in incomes between districts) and between the three large regions (East, Central, West)

China =



Kanbur and Zhang (2003) decomposition has three components:

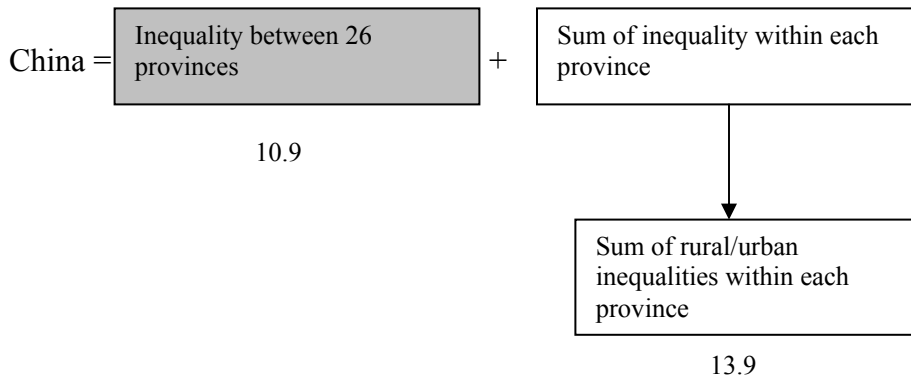


(1) Difference between the overall rural and urban mean, (2) inequality within the urban means (area **U**), and (3) inequality within the rural means (area **R**). The component (1) is equal to 13.9 Theil points, the sum of components (2) and (3) is equal to 10.9 Theil points.

A modified Kanbur-Zhang (K-Zh) decomposition which we use here is to break inequality by provinces. Then, the component (1) is inequality between all provincial (total) means, and the component (2) is the sum of all inequalities between provincial rural and urban means, that is inequality between U_a and R_a , plus inequality between U_b and R_b and so forth. The component (1) amounts to 10.9 Theil points,³⁹ the sum of rural-to-urban inequalities within provinces amounts to 13.9 Theil points.

³⁹ Note that because of Theil's exact decomposition property, inequality among P s must be equal to inequality among U 's plus inequality among R 's, that is the sum of Kanbur and Zhang's components (2) and (3) must be equal to our component (1).

The results of the modified Kanbur and Zhang (K-Zh) decomposition can then be represented as



These two decompositions suggest the following. First, note that the sum of two shaded boxes in the Akira and Kawamura (AK) decomposition equals (by definition) the shaded box in the modified Kanbur and Zhang decomposition.⁴⁰ But while for Akira and Kawamura, the importance of the two shaded boxes adds to 9.4 Theil points, the shaded box in the (modified) Kanbur and Zhang decomposition amounts to 10.9 Theil points. The difference must be due to the use GDP per capita in one study vs. expenditure per capita from household surveys in another.

Second, from the Akira and Kawamura decomposition, we can conclude that most of regional inequality in China is found between the three large regions and within provinces (in other words, provinces contain districts with fairly unequal average incomes). But, the Kanbur and Zhang decomposition suggests that these districts are not random, that is they are not randomly poor or rich. The main line of differentiation goes between rural and urban areas. For while AK decomposition shows that inequality between districts (within all provinces) equals 15.6 Theil points, the K-Zh decomposition

⁴⁰ This must be true for Theil (which both authors use) because of its decomposability.

shows that 13.9 out of these 15.6 Theil points is due to the differences between mean rural and mean urban incomes.⁴¹

Thus, in conclusion, we can say that if we break China into a very fine regional partition, about 40 percent of thus calculated Concept 2 inequality is due to differences in incomes between provinces. The bulk of the remaining 60 percent is due to the differences in mean urban and rural income within provinces.

⁴¹ Abstracting from the fact that the things are not fully comparable because A-K use GDP per capita while K-Zh use household survey data.